









```

11 799 TCGTGGTTCACATATTAATAACCTCCACCTCCCAAGTGGCTCAAGAACCATATAATA 858
13 164 ThrAlaValIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 193
15 859 AAGCGGTGTAACCAATTAAGTGGTTCAGCTGCAACATCCAGCGGTCACCAACAGGAGT 918
17 184 AlaArgIleCatalysIndLeuAspThrArgIleIleGluValLeuCysLysGlu 203
19 919 GACGATATAGCAAAACATATAAATGATACAGAAATATATGAATCTCTGTAAAGAA 978
21 204 HisGlnCysAsnIleAspGluValLysAsnValIlyrPheLysAsnPheIleProPheMet 223
23 979 CATGAAGTAAATAGATAGATGAGTCAAAATGTTATTTCAAAAAATTTATACCTTTATG 1038
25 224 AsnSerLeuLysLeuValThrSerAsnGlyLeuProGluValGluAsnLeuSerLysArg 243
27 1039 AATTCCTTCGACATCTGAACATCTAAGACATCGACAGCTGAAATCTCTTCAACGA 1098
29 244 TyrGluThrIleTyrLeuLysAsnLysAspAlaArgLeuPheLeuAspHisAsp 263
31 1099 TACGAAGAAATTTATCTTAAATAATAAGATCTAGATGCAAGATTATTTTGGATCATGAT 1158
33 264 TyrThrGluGlnThrAspSerIleAspSerPheGluThrGlnArgThrProArgLysSer 283
35 1159 AAAACCTCTTCACACTGATCTATACACAGTTTTCGAACACACAGAACACCAACAAAAGT 1218
37 284 AsnLeuAspGluGlnValAsnValIleProProIleThrProValArgThrValMetAsn 303
39 1219 AACCTGACAGACAGCTGAATGTAATTCCTCCACACACTCCAGTTAGGACGTGTATGAC 1278
41 304 ThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 323
43 1279 ACTATGACAAATTAATGATGATTTTAAATTCACAAATATCAATCAATTCAGAAATCTG 1338
45 324 IleSerTyrPheAsnAspCysThrValAsnProLysGluSerIleLeuLysArgValLys 343
47 1339 ATTCTCTAATTTTAAACACGACAGCTGAAATCCAAACAAAGTATACGAAACAGTGAAG 1398
49 344 AspIleGlyTyrIlePheLysGlnLysPheAlaLysAlaValGlyGlnGlyCysValGlu 363
51 1399 GATAAGATAACATCTTAAACATAAAATTTGCTAAAGCTGGGACACAGGCTTCTGTGAA 1458
53 364 IleGlySerGlnArgThrLysLeuGlyValAlaArgLeuLysArgValMetGluSerMet 383
55 1459 ATGGATACAGATATATAAATTTGATCTTGATTAACCAATTAATGGAATCCATG 1518
57 384 LeuLysSerGlnThrGlnThrLeuSerIleGluAsnPheSerLysLeuLeuAsnAspAsn 403
59 1519 GTTAAATACACACACAGCAATATCCATTCATAAAATTTACCAACATCTGCAATGACAAC 1578
61 404 IlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrTyrSerArg 423
63 1579 ATTTTATAAGTCTTATGCGCTGCGCTCTCGAGGTGTGTAATCGGCACATATAGCAGA 1638
65 424 SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProIleLeuAsnVal 443
67 1639 AHTATATTCGAATCTGATTCGGAAACAGATTGTGTTTGTTCATGATTCGATTCGAATGTG 1698
69 444 LeuAsnLeuLysAlaPheAspPheLysValIleGlnSerPheIleCysArgGlyGly 463
71 1699 GTTAAATTTAAACGCTTCATCTTTACAAAGTGAATCGAAAGTTTATCAAAACGAGAAGCC 1758
73 464 AsnLeuThrArgGluMetIleLysIleLeuGluArgCysGluHisArgIleMetGluSer 483
75 1759 AACCTGATGAAGCAATGATATAAAATTTAGAACCATGCTCAACATCGAAATCATCGAATCC 1818
77 484 LeuAlaIlePheSerAspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGlu 503
79 1819 GTTGAAGCTCTCAATATTCACCTTTATTTATTAACCAATTAATGAAGACCGAGAA 1878
81 504 GlyProThrAspHisLeuGlnSerAlaCysProLeuAsnLeuProLeuGlnAsnHis 523

```

RESULT 3  
AR207831  
LOCUS

2994 bp DNA

AR207831

JAN 20 2003

```

1879 GGAACCACTGATCACTTCAATGCAATGCTTGGCTTAACTTAACTTCTCCACAAATAACAT 1938
524 ThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLysLysGlySerThrThr 543
1939 ATCTGATGATATGATATCTTCTGATTAATGATGATGATGATGATGATGATGATGATGAT 1998
544 ArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLys 563
1999 GGTGTAAATTTATGTAATATGAGAGACAAAGACAAAGCTTCCAGAGACAGAGAG 2058
564 ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyrArgLeuAlaTyrLeu 583
2059 CCATTCGAATCTGCTCTTCTGCTCTTCTGCTCTTCTGCTCTTCTGCTCTTCTGCTCTTCT 2118
584 ArgLeuAsnIleLeuCysGluArgLeuLeuSerGluHisProGluLeuGluHisIleIle 603
2119 GGTGTAAATTTATGTAATATGAGAGACAAAGACAAAGCTTCCAGAGACAGAGAG 2178
604 TyrPheLeuPheGlnHisThrLeuGlnAsnGlnTyrGluLeuMetArgAspArgHisLeu 623
2179 TGTAAATTTATGTAATATGAGAGACAAAGACAAAGCTTCCAGAGACAGAGAG 2238
624 AspGlnIleMetMetCysSerMetTyrGlyIleCysLysValLysAsnIleLeuLeuLys 643
2239 GACCAAAATTTATGTAATATGAGAGACAAAGACAAAGCTTCCAGAGACAGAGAG 2298
644 PheTyrIleIleValThrAlaTyrLysAspLeuProHisAlaValGlnGluThrPheLys 663
2299 TTTAAATTTATGTAATATGAGAGACAAAGACAAAGCTTCCAGAGACAGAGAG 2358
664 ArgValLeuIleLysGluGluGlyTyrAspSerIleIleValPheLysAsnSerValPhe 683
2359 GGTGTAAATTTATGTAATATGAGAGACAAAGACAAAGCTTCCAGAGACAGAGAG 2418
684 MetGlnArgLeuLysThrAsnIleLeuGlnThrAlaSerThrArgProThrLeuSer 703
2419 ATGACACACATCAAAACAAATATTTCCAGATGCTTCCAGAGACAGAGACAGAG 2478
704 ProIleProHisIleProArgSerProTyrLysPheProSerSerProLeuArgIlePro 723
2479 GCAATGCTTATATGCTGAAAGACATTAAGATTTCTATTCATTCATTCATTCATTCATTC 2538
724 CysGlyAsnIleTyrIleSerProGluLysSerThrGlyTyrLysIleSerGluGlyPhePro 743
2539 GAGGATATCTATATTTACCTGCAAGATCAATATAAATTTCCAGAGACAGAGACAGAG 2598
744 ThrProLysMetThrProArgSerArgIleLeuValSerIleGlyGluSerPheGly 763
2599 ACACCAACAAAAACATGCTCCAGATCAAGAAATCTGATCAATTCGATGATTCGATGAT 2658
764 ThrSerGluLysPheGlnLysIleAsnGlnMetValLysAsnSerAspArgValLeuLys 783
2659 ATTTCTGAAATTTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2718
784 ArgSerAlaGluGlySerAspProThrGlyTyrLeuLysLysLeuArgPheAspIleGlu 803
2719 AATAATTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2778
804 GlySerAspGluAlaAspGlySerLysHisLeuProGlyGluSerLysPheGlnLys 823
2779 GCAATGATCAACACAGATCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2838
824 LeuAlaGluMetThrSerThrArgThrArgMetGlnLysIleLysMetAspSerMet 843
2839 CTGGTGAATATGATTTATTCGAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2898
844 AspThrSerAsnLysGluGluLys 851
2899 GATACCTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2922

```



















[illegible]

11 1696 GTTAAATTAAGAGGCTGATGATTTTACAAAGCTGATCCAAAGTTTATCAAGGACAGAGCC 1757  
 12 464 AsnLeuThrArgGluMetIleLysLysIleuArgCysGluLysIleuArgIleMetGluSer 483  
 13 1694 AACTTCATTAAGAGAAATGCAAAATATTTAGACGATGTCGAAATCAATGCAATCC 1817  
 14 483 LeuAlaIleLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGlu 503  
 15 1816 TTGATAGCTCTGATGATGATCTTTATTTGATCTTATTAACCAATCAAGGACCGAGAA 1877  
 16 503 GlyProIleAspHisIleuGlnSerAlaCysProLeuAsnIleuProIleGlnAsnAspHis 523  
 17 1878 GACCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1937  
 18 523 ThrAlaAlaAspMetIleLysSerProValArgSerProLysLysLysLysLysLysLys 543  
 19 1938 ACTGACAGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1997  
 20 543 ArgValAspSerThrAlaAsnAlaIleuThrGlnAlaIleuSerAlaPheGlnThrGlnLys 563  
 21 1998 TGTGATAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2057  
 22 563 ProLeuLysSerThrSerLysSerLeuPheLysLysValIleArgGluAlaIleLys 583  
 23 2057 TGTGATAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2117  
 24 583 ArgLeuAsnThrLeuCysGluArgLeuLeuSerGluLysProGluLeuGluLysIleIle 603  
 25 2117 GATGATAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2177  
 26 603 GlyProIlePheGlnSerThrGlnAlaIleuThrGlnAlaIleuSerAlaPheGlnThrGlnLys 623  
 27 2178 TGTGATAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2237  
 28 623 AspGlnIleMetMetCysSerMetIleCysLysValLysAsnIleAspPheLys 643  
 29 2237 GATGATAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2297  
 30 643 PheLysIleIleValThrAlaIleLysAspLeuProHisAlaValGlnAlaIlePheLys 663  
 31 2298 TTTCAAAATATGCTAACACCAACACCAACACCAACACCAACACCAACACCAACACCAAC 2357  
 32 663 ArgValLeuIleLysThrAsnIleLeuGlnIleuThrAlaSerThrArgProIleLeuSer 703  
 33 2357 TGTGATAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2417  
 34 703 ArgValLeuIleLysThrAsnIleLeuGlnIleuThrAlaSerThrArgProIleLeuSer 723  
 35 2417 ATGACAAATGACAAACAAATGACAAACAAATGACAAACAAATGACAAACAAATGACAA 2477  
 36 723 ProIleProIleIleProArgSerProTyrLysPheProSerSerProLeuArgIlePro 743  
 37 2478 GAAATGACAAATGACAAACAAATGACAAACAAATGACAAACAAATGACAAACAAATGAC 2537  
 38 743 GlyGlyAspIleLysThrSerPheGlnLysSerProTyrLysIleuSerGlnCysPheLys 763  
 39 2537 GAGAGAAAT 2597  
 40 763 ThrProIleLysMetThrArgSerArgIleLeuValSerIleCysLysSerPheLys 783  
 41 2598 ATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 2657  
 42 783 ThrSerGlnLysPheSerThrAlaGlnMetValLysAsnSerArgArgValLysIle 793  
 43 2658 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2717  
 44 793 ArgSerAlaGlnLysSerAsnProIleLysIleuLysLysLeuArgPheAspIleGln 803  
 45 2718 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2777  
 46 803 GlySerAspIleAlaAspIleLysSerLysHisIleuProCysGluSerLysPheClnLys 823

Db 2778 GCATCAGATCAAGCAGATCGAAGTAAATATCTCCAGCAGATCCAAATTCACACAAA 2837  
 QY 824 LeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetAspSerMet 843  
 Db 2838 CTGGGTAAATGACTTCTACTCTAACATCAATATGTAATGTAATGTAATGTAATGTAATG 2897  
 QY 844 AspThrSerAsnLysGluLysLys 851  
 Db 2898 CATATCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2921  
 RESULT 9  
 HUMRH1MRNA  
 LOCUS  
 DEFINITION Homo sapiens retinoblastoma susceptibility protein (Rb1) mRNA and mutations.  
 ACCESSION L41870  
 VERSION L41870.1 GI:794994  
 KEYWORDS retinoblastoma protein; retinoblastoma susceptibility.  
 SOURCE Homo sapiens cDNA to mRNA.  
 ORGANISM Homo sapiens  
 MAMMALIA: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
 REFERENCE 1 (sites)  
 McGeer, P.L., Vandell, P.W. and Dryja, J.P.  
 Structure and partial genomic sequence of the human retinoblastoma susceptibility gene.  
 JOURNAL Gene 80 (1), 119-128 (1989)  
 MEDLINE 90006771  
 PUBMED 2701949  
 REFERENCE 2 (sites)  
 Hogg, A., Onadim, Z., Baird, P.N. and Cowell, J.K.  
 Detection of heterozygous mutations in the Rb1 gene in retinoblastoma patients using single-strand conformation polymorphism analysis and polymerase chain reaction sequence.  
 JOURNAL Oncogene 7 (7), 1445-1451 (1992)  
 MEDLINE 92319557  
 PUBMED 1352398  
 REFERENCE 3 (sites)  
 Onadim, Z., Hogg, A., Baird, P.N. and Cowell, J.K.  
 Genotype-phenotype mutations in exon 30 of the Rb1 gene in families showing incomplete penetrance and mild expression of the retinoblastoma phenotype.  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (13), 6177-6181 (1992)  
 MEDLINE 92335261  
 PUBMED 1352883  
 REFERENCE 4 (sites)  
 Onadim, Z., Hogg, A. and Cowell, J.K.  
 Mechanisms of oncogenesis in patients with familial retinoblastoma.  
 JOURNAL Br. J. Cancer 68 (5), 958-964 (1993)  
 MEDLINE 94031584  
 PUBMED 8217609  
 REFERENCE 5 (sites)  
 Hogg, A., Bial, R., Onadim, Z. and Cowell, J.K.  
 Molecular mechanisms of oncogenic mutations in tumors from patients with bilateral and unilateral retinoblastoma.  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90 (15), 7361-7365 (1993)  
 MEDLINE 93348271  
 PUBMED 8346235  
 REFERENCE 6 (sites)  
 Kratz, R.A., Petersen, G.A., Reed, A., Coxon, A.H., Goratis, J., Cowell, J.K. and Kaye, F.J.  
 Partial inactivation of the Rb product in a family with incomplete penetrance of familial retinoblastoma and benign retinal tumors.  
 JOURNAL Am. J. Hum. Genet. 55 (5), 1321-1326 (1994)  
 MEDLINE 94203660  
 PUBMED 8152792  
 COMMENT Mutation data provided by Dr. R.L. Gallic, Hospital for Sick Children, 555 Univ. Ave., Toronto, ONT M5G 1X8, Canada  
 Ph. 416 813-5868  
 FAX: 416 813-4989

Genature version 5.1.1  
Copyright (c) 1993 - 2003 Computer Ltd.

cm nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:20:22 : Search time 2634.92 seconds  
(without alignments)  
18469.506 Million cell updates/sec

Index: us-09-026-459a-42

Perfect score: 647

Sequence: 1 CGGTATATGAGGAAAAC.....AAATGAGATTAATGATAGT 3347

Scoring table: IDENTITY 90%

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 806774376 residues

Total number of hits satisfying chosen parameters: 32498142

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match ok

Maximum Match 100%

Listing first 45 summaries

Database: EST

- 1: cm-est-hum:\*
- 2: cm-est-hum:\*
- 3: cm-est-hum:\*
- 4: cm-est-hum:\*
- 5: cm-est-hum:\*
- 6: cm-est-hum:\*
- 7: cm-est-hum:\*
- 8: cm-est-hum:\*
- 9: cm-est-hum:\*
- 10: cm-est-hum:\*
- 11: cm-est-hum:\*
- 12: cm-est-hum:\*
- 13: cm-est-hum:\*
- 14: cm-est-hum:\*
- 15: cm-est-hum:\*
- 16: cm-est-hum:\*
- 17: cm-est-hum:\*
- 18: cm-est-hum:\*
- 19: cm-est-hum:\*
- 20: cm-est-hum:\*
- 21: cm-est-hum:\*
- 22: cm-est-hum:\*
- 23: cm-est-hum:\*
- 24: cm-est-hum:\*
- 25: cm-est-hum:\*
- 26: cm-est-hum:\*
- 27: cm-est-hum:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	910.2	27.2	1623	BM459031	BM459031 AGENCOURT
2	909.4	27.2	1051	BM561083	BM561083 AGENCOURT
3	856	25.6	1083	BM439221	BM439221 AGENCOURT
4	844.8	25.2	881	BM220275	BM220275 AGENCOURT
5	762.4	22.5	1122	BM545881	BM545881 AGENCOURT
6	725.4	21.7	727	AL597811	AL597811 DKFZP3130

7	699.6	20.9	823	12	BM610661
8	698.2	20.9	792	9	AL563757
9	692.2	20.7	767	12	BM579456
10	691.8	20.7	922	13	BM466194
11	678	20.3	835	13	BM453724
12	655.2	19.9	1001	13	BM475603
13	660	19.7	672	13	BM186859
14	646.6	19.3	882	14	BM222227
15	641.6	19.2	797	10	BM583278
16	642.2	19.2	581	12	BM514239
17	617	18.4	618	13	BM264009
18	614	18.3	694	13	BM082846
19	605.2	18.1	623	10	BM383181
20	605	18.1	605	9	AL599633
21	596.4	17.8	922	12	BM253543
22	580	17.3	580	9	AL599105
23	576.4	17.2	1364	11	AK011246
24	575.2	17.2	914	9	AA763411
25	566.4	16.9	588	10	AA762234
26	558	16.7	871	9	AA763485
27	530.2	15.8	598	10	AA762967
28	530	15.8	539	9	AL592790
29	528	15.8	572	9	AL130941
30	516.8	15.4	597	12	BM312246
31	513.2	15.3	594	10	AA715533
32	512.2	15.3	734	12	BM144857
33	505.2	15.1	736	13	BM151568
34	493.8	14.8	525	9	AA258255
35	486.4	14.5	584	12	BM149050
36	484.4	14.5	493	10	AA502887
37	481.4	14.4	484	10	BM168095
38	480.2	14.3	659	13	BM121116
39	476.6	14.2	2083	13	BM456728
40	474.4	14.2	734	13	BM152698
41	459.2	13.7	497	12	BM54784
42	457	13.7	457	9	AI093215
43	457	13.7	742	9	AL598766
44	454.2	13.6	869	12	BM65749
45	447	13.4	447	9	AI082179

ALIGNMENTS

RESULT 1  
BM450031  
LOCUS BM450031 1023 bp mpna linear EST OF FEB 2002  
DEFINITION AGENCOURT\_6393544 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5528037  
5' mRNA sequence.  
ACCESSION BM450031  
VERSION BM450031.1 GI:18499071  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 1023)  
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [raupbs@mail.nih.gov](mailto:raupbs@mail.nih.gov)  
Tissue Procurement: AACC/WU/DIP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINC)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINC at:  
<http://image.linc.nih.gov>  
plate: IAM2204 row: d column: 22  
High quality sequence stop: 643.  
Location/Qualifiers  
1. 1023

[illegible]





















Plate: LLAM8427 row: 1 column: 18

High quality sequence stop for

localization effects

FLA0015

Source

1: 797

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3449873"

/clone\_lib="NIH\_MGC\_10"

/cell\_line="MGC36"

/lab\_host="DH10B"

/note="organ: corvix; Vector: pCMV-Sport6; Site: 1; Not1;

Site 2; Salt: Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.5 kb. Library prepared by Life

Technologies."

Blast: 0.01 252 a 171 c 167 q 207 t

K101N

Query Match 19.2% Score 643.6; DB 10; Length 797;

Best Local Similarity 95.8% Pred. No. 2.2e-115;

Matches 747; Conservative 0; Mismatches 24; Indels 9; Gaps 8;

QY 2000 CTGCTCAGCAGACATTCACACGCTTTTCATCAACAGACAGACAGCTATGATCTCTATTATAG 2059

DB 1 CTGCTCAGCAGACATTCACACGCTTTTCATCAACAGACAGACAGCTATGATCTCTATTATAG 59

QY 2060 TATTTCTATAACTCGCTCTTCTATGACGASACTGAAACAAATATTTTGTGATGTCTTCCA 2119

DB 60 TATTTCTATAACTCGCTCTTCTATGACGASACTGAAACAAATATTTTGTGATGTCTTCCA 118

QY 2120 GCAGAGCGGCTATGTTGTGACCAATAGCTACATTCCTGCGAGAGGCTTTACAGTTTCTTA 2179

DB 119 GCAGAGCGGCTATGTTGTGACCAATAGCTACATTCCTGCGAGAGGCTTTACAGTTTCTTA 178

QY 2180 GTTCATGCTTACGCAATCTCCAGAGGCAACATCTATATTCACCCCGGSAAGAGTCCATATA 2239

DB 179 GTTCATGCTTACGCAATCTCCAGAGGCAACATCTATATTCACCCCGTGAAGAGTCCATATA 238

QY 2240 AAATTTTAAAGAGCTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2299

DB 239 AAATTTTAAAGAGCTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 298

QY 2400 CAATTTGTGAATTCATCTCGGAGATTTGTGAGAAATTCAGAAATTAATTAATTAATTAATTAAT 2459

DB 299 CAATTTGTGAATTCATCTCGGAGATTTGTGAGAAATTAATTAATTAATTAATTAATTAATTAAT 358

QY 2460 AATGAAATTTGCTTCAAAAGAAATGCTTCAAGGAAAGAA-CCCTCCCTAAACCCACTGAAA 2418

DB 459 AATGAAATTTGCTTCAAAAGAAATGCTTCAAGGAAAGAA-CCCTCCCTAAACCCACTGAAA 418

QY 2419 AAATCTAGCTTTTCTATATTCAGAGATCTCATATGAGAGATGGAAGTAAATCATCTCTCCAGGA 2478

DB 419 AAATCTAGCTTTTCTATATTCAGAGATCTCATATGAGAGATGGAAGTAAATCATCTCTCCAGGA 478

QY 2479 GATTTCAAAATTTCTGAAATTTGCTTCAAGGAAAGAA-CCCTCCCTAAACCCACTGAAA 2537

DB 479 GATTTCAAAATTTCTGAAATTTGCTTCAAGGAAAGAA-CCCTCCCTAAACCCACTGAAA 538

QY 2538 GCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2597

DB 539 GCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 598

QY 2598 CTTTGTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2657

DB 657 CTTTGTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 714

QY 2718 AATGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2777

DB 773 AATGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 773

Search completed January 18, 2003, 04:09:46  
Job time: 2962.99 secs



Core version 5.1.4  
Copyright © 1993 - 2003 CompuGen Ltd.

on protein-nucleic search, using frame\_plus\_p2n model

Run on: January 17, 2003, 18:27:24 ; Search time 1840.65 Seconds  
(without alignments)  
7320.587 Million cell updates/sec

115 09 026 459A 45

1901 1001 55010 : 4278

Sequence: 1 MLW; 1' FIAAVDLDLDEMSPT ... TMQKQKMNISMDSNKEEK 832

NOTES FOR THE EDITOR:  $\text{Hf}_2\text{SiMg}_2$

<b>Xq1p0p</b>	10, 0	<b>Xq1p0xt</b>	0, 5
<b>Yq1p0p</b>	10, 0	<b>Yq1p0xt</b>	0, 5
<b>Fq1p0p</b>	6, 0	<b>Fq1p0xt</b>	7, 0
<b>10c10p</b>	6, 0	<b>10c10xt</b>	7, 0

16154066 seeds, 8097743376 residues

[illegible]

William [sic] Lee

[illegible]

Post process: Minimum  $M_{\text{star}} = 0.8$  ( $M_{\odot}$ )

Maximum Moisture 100%

## Listing first 45 summaries

### Recommended Reading:

[illegible]

1:	em_est_ba:*
2:	em_est_hum:*
3:	em_est_lo:*
4:	em_est_mu:*
5:	em_est_mv:*
6:	em_est_pl:*
7:	em_est_ro:*
8:	em_est_v:*
9:	qb_est_1:*
10:	qb_est_2:*
11:	qb_est_3:*
12:	qb_est_4:*
13:	qb_est_5:*
14:	em_est_1_u:*
15:	em_est_2_u:*
16:	em_est_3_u:*
17:	qb_oss:*
18:	em_oss_hum:*
19:	em_oss_lov:*
20:	em_oss_pl:*
21:	em_oss_vt:*
22:	em_oss_mu:*
23:	em_oss_mv:*
24:	em_oss_nam:*
25:	em_oss_mus:*
26:	em_oss_lov:*
27:	em_oss_pl:*
28:	em_oss_rov:*

## SUMMARY

Result No.	Score	Query Match	Length	FB	IP	Project
1	1551	46.3	1051	13	BM561083	BM561083 AGEN000R1
2	1473	34.4	1364	11	AK011246	AK011246 MOS_MOUSE0
3	1448	33.8	881	14	BQ220275	BQ220275 AGEN000R1
4	1339	31.3	1001	13	BM475603	BM475603 AGEN000R1
5	1254	29.3	727	9	AL159781	AL159781 UKT2P3140
6	1236.5	28.9	883	14	BQ220277	BQ220277 AGEN000R1
7	1208	28.2	835	11	HM453724	HM453724 AGEN000R1
8	1111	26.0	1023	13	BM500041	BM500041 AGEN000R1
C 9	1056.5	24.7	914	9	AA764341	AA764341 wv54304.1
C 10	1043	24.4	792	9	AL563757	AL563757 AL563757
C 11	1037	24.2	605	9	AL599633	AL599633 UKT2P314P
C 12	1030	24.1	694	10	BQ220246	BQ220246 MO2_H1064
C 13	1016	23.7	623	10	AA583181	AA583181 ial11009.1Y
C 14	1013	23.7	871	9	AA763445	AA763445 wv54101.1Y
C 15	1004	23.5	922	12	BQ220243	BQ220243 602364278
16	978	22.9	588	10	AA762834	AA762834 CM3_H1011B
17	962	22.5	736	13	BL151568	BL151568 602921610H
18	959	22.4	594	10	AA719050	AA719050 AL5715533
19	955	22.3	584	12	BQ149050	BQ149050 m088601.1Y
20	947.5	22.1	734	12	BF144857	BF144857 601790443
21	935	21.9	539	9	AL1692705	AL1692705 UKT2P3131
22	908	21.2	737	10	BF539278	BF539278 601063567
23	882	20.6	572	9	AL125943	AL125943 UKT2P312D
24	879.5	20.6	734	13	BL125982	BL125982 602327177
25	869.5	20.3	625	12	BQ241088	BQ241088 602358151
C 26	858	20.1	597	12	BF332346	BF332346 MR0_H1015
C 27	844	19.7	484	10	BE168095	BE168095 QV3_H1051
C 28	843.5	19.5	646	13	BM490316	BM490316 f0p2n_pk0
C 29	824	19.3	493	10	AA502887	AA502887 U1_HF-BN0
30	815.5	19.1	522	9	AA072789	AA072789 mm76a05.1Y
31	784	18.3	2083	13	BM456728	BM456728 AGEN000R1
32	757	17.6	453	10	AA503514	AA503514 U1_HF-BN0
33	721	17.0	832	9	AL164608	AL164608 wv54310.1X
C 34	683	16.0	438	12	BF27047	BF27047 MO3_HN002
C 35	681.5	15.9	799	14	CB2304	CB2304 CB2304 Leuk
36	667	15.6	651	13	BL169684	BL169684 603344967
C 37	659	15.4	380	10	AA156104	AA156104 ES1363674
C 38	658	15.4	462	9	AL1066529	AL1066529 v027q08.1Y
C 39	633	14.8	425	10	BE285796	BE285796 601096510
C 40	628	14.7	672	13	BL186459	BL186459 603393230
C 41	626	14.6	521	12	BF466536	BF466536 U1_M_C00P
C 42	625.5	14.6	387	12	BF998493	BF998493 CM2_GN016
C 43	613	14.3	415	9	AL1745644	AL1745644 Tr2306.X
C 44	611	14.3	410	12	BF560103	BF560103 U1_R-AT1
C 45	609	14.2	411	9	AL1745643	AL1745643 Tr2309.X

## ALLEGMENTS

RESULT 1	
BW561083	
LOCUS	1051 bp mRNA linear EST_20 FEB-2002
DEFINITION	BW561083 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:595065
ACCESSION	5' mRNA Sequence.
VERSION	BW561083
KEYWORDS	HW561083.1 GI:18806043
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo,
REFERENCE	1 (bases 1 to 1051)
AUTHORS	NIH-Mac <a href="http://mac.ncbi.nlm.gov/">http://mac.ncbi.nlm.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)











[illegible]









Gulimova, G. H., Caralho, A. J., Matsukuma, A., Bahia, S., Simpson, D. H., Brangulienė, A., de Oliveira, P. S., Bucher, P., Jouque, J. V., O'Hare, M. J., Soares, E., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J.

Shortgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 4491-4496 (2000)

20020664

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, São Paulo SP, Brazil.

tel: +55 11 2704922

fax: +55 11 2707001

email: asimpson@ludwig.org.br

This sequence was derived from the FAPES/Ludwig human Cancer Genome Project. This entry can be seen in the following URL:

(<http://www.ludwig.org.br/cg/submit.cgi?A=2-P-2-RI-642-230>)

2003014-009683-2000-02-21614-1)

Seq primer: puc 18 forward

High quality sequence start: 58

High quality sequence stop: 669.

Location/Qualifiers

1..694

/organism "Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="H10642"

/dev\_stage="Adult"

/note="organ: breast; vector: puc18; site: 1; small; site-2;

Small: A mutant library was made by cloning products derived

from ORFESTES PCR (O.S. Letters Patent application No. 196

2716 - Ludwig Institute for Cancer Research) profiles

into the puc 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

BASE COUNT 211 a 129 c 127 g 226 t 1 others

ORIGIN

#### Alignment Scores:

Prod. No.: 1..694 100 Length: 694  
Score: 1040.00 Matches: 207  
Percent Similarity: 96.44% Conservative: 3  
Best Local Similarity: 94.95% Mismatches: 7  
Query Match: 24.06% Indels: 2  
Gaps: 0

75 99 926 459a 45 (1 832) x BE082846 (1-694)

27 410 AcgTysThrValAsnPrdGlyGluSerPheLeuLysArgValLysAcgTleClyTyrTle 429

146 676 CACGACACAGAC CAAAAGAAAGAGACACGACAGACAGAGAGAGACAGACAC 618

27 440 PheGlySerGlySerAlaValGlyGlnGlyGlyValGlyGlnGlySerGlnArg 449

146 617 TTTAAAGAGAAATTGTGTAAGCTGTGGACAGGTTGTGTGCAAAATGATATCAAGCGA 558

27 450 TTTTysLeuGlyValAlaTleClyTyrAlaValMetGluSerMetLeuLysSerGlnGlu 369

146 647 TAAATTTGCAATTCGTTGTATATACGATATAGGATATGATGTTTAAATCAAGAAG 498

27 470 GTCACLeuSerPheGluAsnPheSerLysLeuLeuAsnAsnLlePheHisMetSer 489

146 497 CAAGATTAATGATATCAAAATTAAGAAATTCGAAATCAACATTTTCATATGCTC 438

27 490 LeuLeuAlaLysAlaLeuGluValMetAlaThrTyrSerArgSerThrSerGlnAsn 409

146 447 TTTATGGAGTGGCTCTTAAAGATGAGATATAGAGATATATATGATATGATGAT 478

27 410 LeuAspSerGlyThrAspLeuSerPheSerPheLeuLeuAsnValLeuAsnLeuLysAla 429

146 477 CTGATTTGGAAAGATATATCTTTTCATGATGATGATGATGATGATGATGATGATGAT 418

QY 440 PheAspPheTyrLysValIleGluSerPheIleLysAlaValGlyGlnGlySerGlnArg 449  
DB TTTTCAATTTTACAAAGCATCGAAAGATTTTATCAAGCAAGCAAGCAAGCAAGCAAGCA 258  
QY 450 MetIleLysHisLeuGluAlaGlyCysGlnHisAlaIleMetGluSerLeuAlaIlePheSer 469  
DB ATGATAAATATTAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198  
QY 470 AspSerPheLeuPheAspLeuIleLysLysLysLysLysLysLysLysLysLysLysLys 489  
DB GATTCACCTTATTTGATGATTTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 138  
QY 490 LeuGluSerAlaCysProLeuAsnLeuProLeuGluAsnAsnHisThrAlaAlaAspMet 509  
DB CTTCAAATCTGTTGCTGCTTAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 78  
QY 509 TTTTLeuSerProValArgSerProLysLysLysLysLysLysLysLysLysLysLysLys 526  
11 77 TATTTTCTGTAATATTTTAAAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTT 26

#### FEATURES

AW583181

IMAGE:5648049 5' similar to gb:M15400 PFTINORLAPSTOMA cDNA clone

PROTEIN (HUMAN): gb:M26491 Mouse reticulostoma susceptibility

PROTEIN (MUSK): gb:M26491 Mouse reticulostoma susceptibility

AW583181

EST:

AW583181.1 GI:7260125

LOCUS

DEFINITION

624 bp mRNA linear 181 19 MAR 2002

IMAGE:5648049 5' similar to gb:M15400 PFTINORLAPSTOMA cDNA clone

PROTEIN (HUMAN): gb:M26491 Mouse reticulostoma susceptibility

PROTEIN (MUSK): gb:M26491 Mouse reticulostoma susceptibility

AW583181

EST:

AW583181.1 GI:7260125

LOCUS

DEFINITION

624 bp mRNA linear 181 19 MAR 2002

IMAGE:5648049 5' similar to gb:M15400 PFTINORLAPSTOMA cDNA clone

PROTEIN (HUMAN): gb:M26491 Mouse reticulostoma susceptibility

PROTEIN (MUSK): gb:M26491 Mouse reticulostoma susceptibility

AW583181

EST:

AW583181.1 GI:7260125

LOCUS

DEFINITION

624 bp mRNA linear 181 19 MAR 2002

IMAGE:5648049 5' similar to gb:M15400 PFTINORLAPSTOMA cDNA clone

PROTEIN (HUMAN): gb:M26491 Mouse reticulostoma susceptibility

PROTEIN (MUSK): gb:M26491 Mouse reticulostoma susceptibility

AW583181

EST:

AW583181.1 GI:7260125

LOCUS

DEFINITION

624 bp mRNA linear 181 19 MAR 2002

IMAGE:5648049 5' similar to gb:M15400 PFTINORLAPSTOMA cDNA clone

PROTEIN (HUMAN): gb:M26491 Mouse reticulostoma susceptibility

PROTEIN (MUSK): gb:M26491 Mouse reticulostoma susceptibility

AW583181

EST:

AW583181.1 GI:7260125

LOCUS

DEFINITION

624 bp mRNA linear 181 19 MAR 2002

IMAGE:5648049 5' similar to gb:M15400 PFTINORLAPSTOMA cDNA clone

PROTEIN (HUMAN): gb:M26491 Mouse reticulostoma susceptibility

PROTEIN (MUSK): gb:M26491 Mouse reticulostoma susceptibility

AW583181

EST:

AW583181.1 GI:7260125

LOCUS

DEFINITION

624 bp mRNA linear 181 19 MAR 2002

IMAGE:5648049 5' similar to gb:M15400 PFTINORLAPSTOMA cDNA clone

PROTEIN (HUMAN): gb:M26491 Mouse reticulostoma susceptibility

PROTEIN (MUSK): gb:M26491 Mouse reticulostoma susceptibility

AW583181

EST:

AW583181.1 GI:7260125

BASE COUNT 244 a 94 c 109 g 186 t 1 others

tel: 617-495-1812

fax: 617-495-8557

Email: dmellon@chicp.harvard.edu

Libraries were constructed by Dr. Douglas Melton

DNA sequencing by: Washington University Genome Sequencing Center

For information on obtaining a clone please contact: Juliana Brown

(brown@fas.harvard.edu)

This sequence now available from the IMAGE consortium, for clone

orders contact: info@image.lln.gov

Seq primer: -406p from Gibco

High quality sequence stop: 406.

Location/Qualifiers

1..624

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="IMAGE:5648049"

/clone\_lib="Human Pancreatic Islets"

/tissue\_type="Islets of Langerhans"

/dev\_stage="Adult"

/lab\_host="pBil08"

/note="organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1;

Site 2: Sal 1; Library constructed using SuperScript

plasmid library kit (Life Technologies); cDNA made by

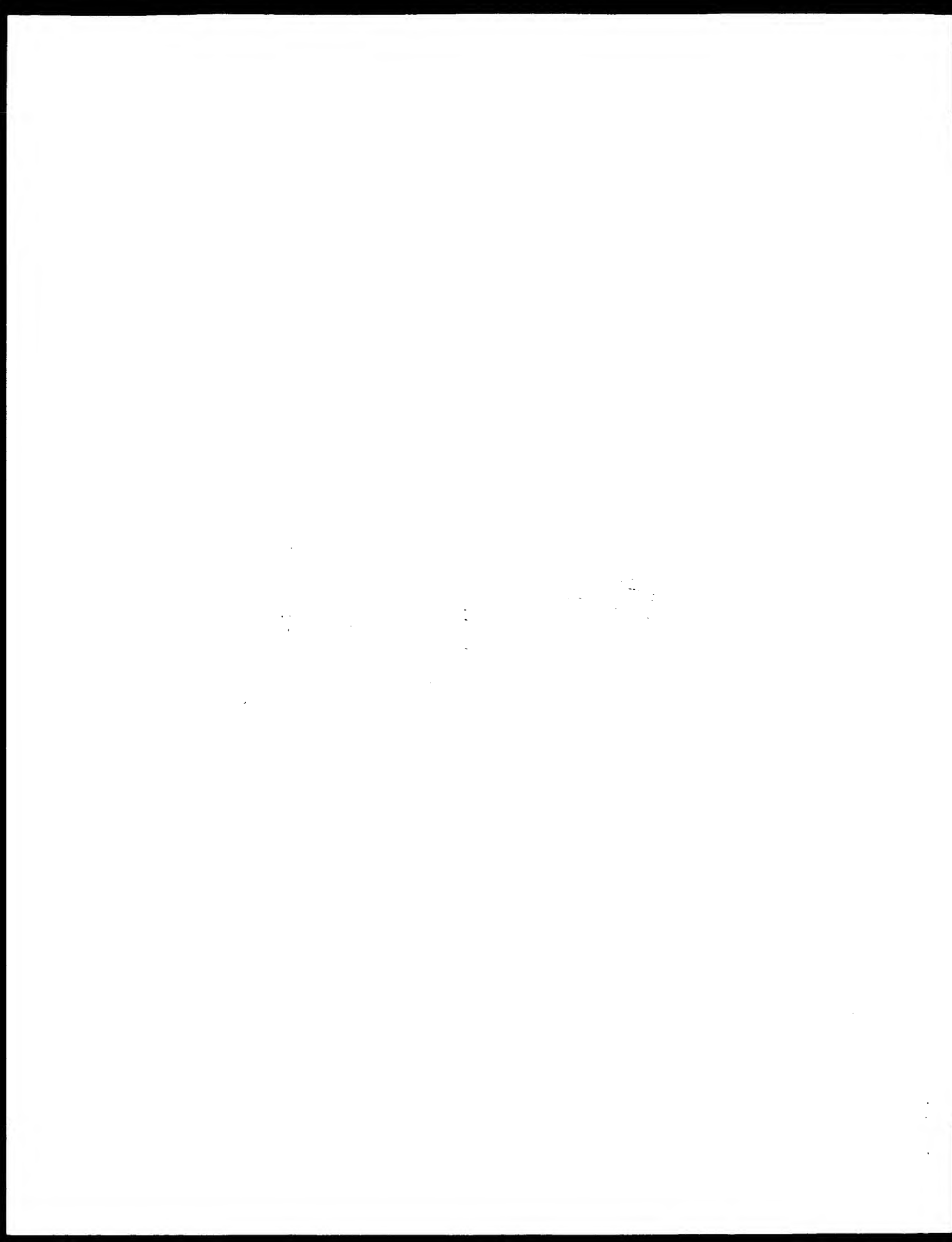
oligo dI priming. Size selected by column fractionation;

average insert size 1.08 kb. Primary library.

unamplified."







megacore version 5.1.4  
Copyright (c) 1994 - 2003 Computer Ltd.

OM nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:20:22 : Search time 2729.72 Seconds  
(without alignment)  
18469,506 Million cell updates/sec

Table: us-09-026-459a-36

Reflected score: 4114  
Sequence: 1 GGGGAGAG-AGAG-AGG.....AAAGAGAGAGAGAGAGT 3113

Scoring table: IDENTITY\_NP

Gap: 10.0 : Gapext 1.0

Search: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 3239432

Minimum db seq length: 0

Maximum db seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST :

- 1: em\_est1a.\*
- 2: em\_est1b.\*
- 3: em\_est1c.\*
- 4: em\_est1d.\*
- 5: em\_est1e.\*
- 6: em\_est1f.\*
- 7: em\_est1g.\*
- 8: em\_est1h.\*
- 9: em\_est1i.\*
- 10: em\_est1j.\*
- 11: em\_est1k.\*
- 12: em\_est1l.\*
- 13: em\_est1m.\*
- 14: em\_est1n.\*
- 15: em\_est1o.\*
- 16: em\_est1p.\*
- 17: em\_est1q.\*
- 18: em\_est1r.\*
- 19: em\_est1s.\*
- 20: em\_est1t.\*
- 21: em\_est1u.\*
- 22: em\_est1v.\*
- 23: em\_est1w.\*
- 24: em\_est1x.\*
- 25: em\_est1y.\*
- 26: em\_est1z.\*
- 27: em\_est1aa.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length (b)	ID	Description
1	910.2	29.2	1023	13	BM450031 AGENCOURT
2	909.4	29.2	1023	14	BM561083 AGENCOURT
3	856	27.5	1081	14	BM439221 AGENCOURT
4	844.8	27.1	881	14	BM228275 AGENCOURT
5	793.6	24.2	1001	13	BM475603 AGENCOURT
6	752.4	24.2	1122	13	BM545881 AGENCOURT

7	725.4	23.3	727	9	AL597811
8	720.6	23.1	880	14	BM222227
9	699.6	22.5	823	12	BM610661
10	698.2	22.4	792	9	AL564757
11	692.2	22.2	767	12	BM570456
12	691.8	22.2	972	13	BM466194
13	678	21.8	835	13	BM453724
14	665	21.2	672	13	BM68355
15	650.2	20.9	1464	11	AK011246
16	643.6	20.7	797	10	HE549278
17	642.2	20.6	681	12	BM616219
18	617	19.8	618	13	BM264009
19	614	19.7	694	10	BM082846
20	605.2	19.4	623	10	AM584181
21	605	19.4	605	9	AL599633
22	596.4	19.2	922	12	BM253533
23	580	18.6	580	9	AL596105
24	575.2	18.5	914	9	AM764811
25	566.4	18.2	588	10	AM668234
26	558	17.9	871	9	AM76485
27	530.2	17.0	598	10	AM582957
28	530	17.0	539	9	AL692790
29	528	17.0	572	9	AL128941
30	516.8	16.6	597	12	HE941246
31	513.2	16.5	594	10	AM715533
32	512.2	16.2	734	12	BE144857
33	505.2	16.2	736	13	RI151568
34	493.8	15.9	525	9	AA258255
35	486.4	15.6	584	12	BM149050
36	484.4	15.6	493	10	AM502887
37	481.4	15.5	484	10	BE168095
38	480.2	15.4	659	13	BM211116
39	476.6	15.3	2083	13	BM456728
40	474.4	15.2	734	13	RI26982
41	459.2	14.8	497	12	HE94784
42	457	14.7	457	9	AL094215
43	457	14.7	732	9	AL598766
44	454.2	14.6	869	12	HE65749
45	447	14.4	447	9	AM082179

ALIGNMENTS

RESULT 1	BM450031	BM450031	6393544	NTH_MGC_72	Homo sapiens cDNA clone IMAGE:5528037	Linear	EST 05 FEB 2002
LOCUS	AGENCOURT	5' mRNA sequence	1023 bp	mRNA			
DEFINITION	BM450031	BM450031.1	G1:18499071				
ACCESSION	BM450031	BM450031.1	G1:18499071				
VERSION	BM450031	BM450031.1	G1:18499071				
KEYWORDS	EST	human					
SOURCE	EST	human					
ORGANISM	EST	human					
REFERENCE	EST	human					
AUTHORS	EST	human					
TITLE	EST	human					
JOURNAL	EST	human					
COMMENT	EST	human					

Human sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 1023)  
NIH-MGC <http://imgc.ncbi.nlm.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [raeborg@mail.nih.gov](mailto:raeborg@mail.nih.gov)  
Tissue Procurement: ATCC/DC/ID/DP  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone Distribution by: Agencourt Bioscience Corporation  
Found through the I.M.A.G.E. Consortium/LLNL at:  
<http://imgc.ncbi.nlm.nih.gov>  
Plate: LAM2204 Row: d Column: 22  
High quality sequence stop: 643  
Location/Qualifiers  
source 1..1023

	QY	2639	GTTATTATTAACCATTCCTGTAAAGTGGTGATGCGTAGCCTGATTTAAAAGTGTGAATAAAGA	2598
	Dg	841	CCTATTATATCAATATGAAAAATCTGTGTAATCCGCCATTAATAAAGATGATGCACA	940
	QY	2599	TTCCTT-CCTCTCCAAGTAAAAATTCCTGTGCTTATAGCAACTAACCAATCC---CCCC	2654
	Ld	601	TT	
	QY	2655	AAGCTGGCACCTCCCA:AAGCACAACCCTCTCTCAGCTACTTCCTCTCTCTCTCTCTCTCT	2714
	Dg	961	AAGCGGCACCTCTCAATATCCCACAGGCTGTCTTCAGTACTTCTGTCTCTCTCTCTCTCT	1020
	QY	2715	A 2715	
	Ld	1021	A 1021	
RESULT 2 BM561083 LOCUS	DEFINITION	RM561083 1051 bp mPRA linear EST 20 FEB 2002		
		AGNCOURTL_656133 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5560905		
		5' mRNA sequence.		
	ACCESSION	BM561083		
	VERSION	BM561083.1 GI:18806043		
	KEYWORDS	EST		
	SOURCE	human		
	ORGANISM	Homo sapiens		
	REFERENCE	Nukaya et al; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1051)		
	AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGSC) Unpublished (1999)		
	JOURNAL	Contact: Robert Strausberg, Ph.D. Email: ccapbes@remail.nih.gov Tissue Procurement: ATCC		
	COMMENT	cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) Cloning Distribution by: Agencourt Bioscience Corporation Clone distribution: MGC clone identification information can be found through the I.M.A.G.E. Consortium/LIML at: <a href="http://image.llnl.gov/">http://image.llnl.gov/</a> Plate: LIML2263 row: n column: 18 High quality sequence stop: 737. Location/Qualifiers 1..1051 organism="Homo sapiens" db_xref="taxon:9606" clone_image=5550905 cdna_lib="NHLI_MGC_67" tissue_type="retinoblastoma" lab_host="PHIOB (phage resistant)" note="Organ: eye; Vector: pCMV-Sport6; Site: 1; Note: Site 2, SalI; Cloned constitutionally. Primer: oligo dT." Average insert size 1.75 kb. Library constructed by Life Technologies."		
	FEATURES	BASIC COUNT 330 A 221 C 197 G 299 T 4 others		
	SOURCE	ORIGIN		
		Query Match                  29.4%      Score 909.4; DB 13; Length 1051;		
		Best Local Similarity       98.4%; Pred. No. 4, 9e-165;		
		Motif's num; Conservation     9; Mismatch'es 13; Indels    2; gaps    2;		
	QY	850	AAAGCTGGCGCACAGGGTTGGCGCGGAATGGAATCACAGCGATACAAACTTTGGAGTTGCG	918
	Dg	1	ATAGCTGTGGACAGAGTTGT	60
	QY	919	TGTATATTTAGTAATTSAAAT	978
	Dg	61	TCTAT	120
	QY	979	AATTTTACCAAACCTCTCTSAATGACAA/TATTTTATATGCTTTTATTGCGGAGCTCT	1048



```

10 121 AATTTTAAATTTTCTGAAATGACAAATTTTTTCAATGCTGCTGCTGCTGCT 180
10 109 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1098
10 181 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 240
10 1099 TGGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1158
10 241 TGGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 300
10 1159 ATGCAAAATTTTATCAAAATGAGCAATATAGCAATATAGCAATATAGCAATATAG 1218
10 301 ATGCAAAATTTTATCAAAATGAGCAATATAGCAATATAGCAATATAGCAATATAG 360
10 1219 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1278
10 461 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 420
10 1279 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1338
10 421 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 480
10 1339 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1398
10 481 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 540
10 1399 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1458
10 541 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 600
10 1459 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1518
10 601 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 660
10 1519 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1578
10 661 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 720
10 1579 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1638
10 721 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 780
10 1639 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1698
10 781 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 840
10 1699 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1757
10 841 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 900
10 1759 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 1809
10 901 GAGTGTGTAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 953

RESULTS 4
LOCUS B04.49221 1081 bp mRNA linear EST 24 MAY 2002
DEFINITION AGC030001_2765562 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015480
5' mRNA sequence.
ACCESSION B04.49221
VERSION B04.49221.3 G1:2178297
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1081)
NIH_MGC ht tp://mgi.nlm.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Published (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: egapbs@mail.nih.gov

```

tissue procurement: AACC  
 cDNA library prepared by: Life Technologies, Inc.  
 DNA Library Arrayed by: The L.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Amersham Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the L.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLNL3211 row: k column: 01  
 High quality sequence stop: 742.

# FEATURES

## source

Location/Qualifiers  
 1..1081  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6015480"  
 /clone\_lib="NIH\_MGC\_92"  
 /issue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="organism: testis; vector: pCMV-Sp-R1; Site 1: Not1;  
 Site 2: SalI; Cloned unidirectionally; original primed.  
 Average insert size: 2.5 kb. Library enriched for  
 full length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

BASE COUNT 437 a 190 c 183 g 371 t

Query Match 27.5% Score 856; Db 14; Length 1081;  
 Best Local Similarity 99.1% Pred. No. 9, w 199;  
 Matches 892; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

QY 2215 GAAGCAGATGCAAGCAAAATCTCCACAGACAGATCCAAATTTTACGACAAATTTGACAA 2274  
 DB 1 GAGACAGATGCAAGCAAAATCTCCACAGACAGATCCAAATTTTACGACAAATTTGACAA 60  
 QY 2275 ATGACTTCTACTCGAAGCAAGCAAAATCTCCACAGACAGATCCAAATTTTACGACAA 2334  
 DB 61 ATGACTTCTACTCGAAGCAAGCAAAATCTCCACAGACAGATCCAAATTTTACGACAA 120  
 QY 2335 AACAGCAAGCAAGCAAAATCTCCACAGACAGATCCAAATTTTACGACAAATTTGACAA 2394  
 DB 121 AACAGCAAGCAAGCAAAATCTCCACAGACAGATCCAAATTTTACGACAAATTTGACAA 180  
 QY 2395 CATCT 2454  
 DB 181 CATCT 240  
 QY 2455 ATATCT 2514  
 DB 241 ATATCT 300  
 QY 2515 AAGCCATCT 2574  
 DB 301 AAGCCATCT 360  
 QY 2575 TGGCATTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 2634  
 DB 361 TGGCATTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 420  
 QY 2635 TGGCATTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 2694  
 DB 421 TGGCATTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 480  
 QY 2695 TGGCATTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 2754  
 DB 481 TGGCATTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 540  
 QY 2755 TTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 2814  
 DB 541 TTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 600  
 QY 2815 TTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 2874  
 DB 601 TTTTAAATGAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAGCAATATAG 660

```

2875 1AGAAATACAAAAAATACAAATTTTACACATTAGATTTTATTTACTTATGCAATCT 2934
|||||
661 1AGAAATAG-AAAAATTACTAAATTTTACACATTAGATTTTATTTACTTATGCAATCT 719
|||||
2945 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 2994
|||||
720 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 778
|||||
2995 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 3054
|||||
779 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 838
|||||
3055 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 3113
|||||
839 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 898
|||||

RESULT 4
BQ220275
DEFINITION AGENT: BQ220275 Homo sapiens cDNA clone IMAGE:6044522
5' mRNA sequence.
ACCESSION BQ220275
VERSION BQ220275.1 GI:20401675
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
1 (bases 1 to 881)
NIH-MGC http://mgc.nhl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: eqapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Cloned distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13287 row: 6 column: 03
High quality sequence stop: 690.
Location/Qualifiers
1..881
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MGC-6044522"
/clone_lib="NIH MGC 92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="BHK21 (phage-resistant)"
/notes="Original test is: Vector; pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size: 2.5 kb. Library method for
full-length clones and constructed by Life Technologies
Note: this is a NIH-MGC library."

BASE COUNT 414 a 149 g 155 q 264 t
ORIGIN

Query Match 27.1% Score 844.8; DB 14; Length 881;
Best local Similarity 99.0%; Pred. No. 130-152;
Matches with conservation 7; Indels 2; Gaps 2.

2875 1AGAAATACAAAAAATACAAATTTTACACATTAGATTTTATTTACTTATGCAATCT 2934
|||||
661 1AGAAATAG-AAAAATTACTAAATTTTACACATTAGATTTTATTTACTTATGCAATCT 719
|||||
2945 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 2994
|||||
720 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 778
|||||
2995 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 3054
|||||
779 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 838
|||||
3055 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 3113
|||||
839 1GATATATTTGCTGCTGTTTATATAAAATTTTCTTTTAAATTAATAAAGGTTGGAAGCAA 898
|||||

RESULT 5
BQ475603
DEFINITION AGENT: BQ475603 NIH-MGC_92 Homo sapiens cDNA clone IMAGE:5575683
5' mRNA sequence.
ACCESSION BQ475603
VERSION BQ475603.1 GI:18524645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
1 (bases 1 to 1001)
NIH-MGC http://mgc.nhl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: eqapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

```

|||||
121 ATGTTTCACTTCGACACACACACACACACACACACACACACACACACACACACAC 180
|||||
365 TAGAAATATGATACAGAAATATTGAAGTTCTCTGTAAGAACATGTAATGTAATGATG 424
|||||
181 TAGAAATATGATACAGAAATATTGAAGTTCTCTGTAAGAACATGTAATGTAATGATG 240
|||||
425 AGTGAAGAAATGTTTATTTCAGAAATTTTATACCTTTTATCAATTTCTTGACATGAA 484
|||||
241 AGTGAAGAAATGTTTATTTCAGAAATTTTATACCTTTTATCAATTTCTTGACATGAA 500
|||||
485 CATCTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 544
|||||
301 CATCTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 460
|||||
545 AATATTAACATTTAGATGTCAGATTTATTTTGGATCATGATAAACTCTTCAACTGAT 604
|||||
361 AATATTAACATTTAGATGTCAGATTTATTTTGGATCATGATAAACTCTTCAACTGAT 420
|||||
605 CTAATACATCTTTCAGAAATATACACACACACACACACACACACACACACACAC 664
|||||
421 CTAATACATCTTTCAGAAATATACACACACACACACACACACACACACACACAC 480
|||||
665 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 724
|||||
481 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 540
|||||
725 TATTTTAAATTCAGTAAATGATACATCTTCTTAAACGATACCAAGAAATATTA 784
|||||
541 TATTTTAAATTCAGTAAATGATACATCTTCTTAAACGATACCAAGAAATATTA 600
|||||
785 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 844
|||||
601 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 660
|||||
845 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 904
|||||
661 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 720
|||||
905 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 964
|||||
721 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 780
|||||
965 GATTATCCATTCAGTAAATGATACATCTTCTTAAACGATACCAAGAAATATTA 1022
|||||
781 GATTATCCATTCAGTAAATGATACATCTTCTTAAACGATACCAAGAAATATTA 840
|||||
1023 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 1080
|||||
841 ATGTAATGCACTTCACAGAGTTGAAATATCTTCTTAAACGATACCAAGAAATATTA 880
|||||

RESULT 5
BQ475603
DEFINITION AGENT: BQ475603 NIH-MGC_92 Homo sapiens cDNA clone IMAGE:5575683
5' mRNA sequence.
ACCESSION BQ475603
VERSION BQ475603.1 GI:18524645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
1 (bases 1 to 1001)
NIH-MGC http://mgc.nhl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: eqapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```











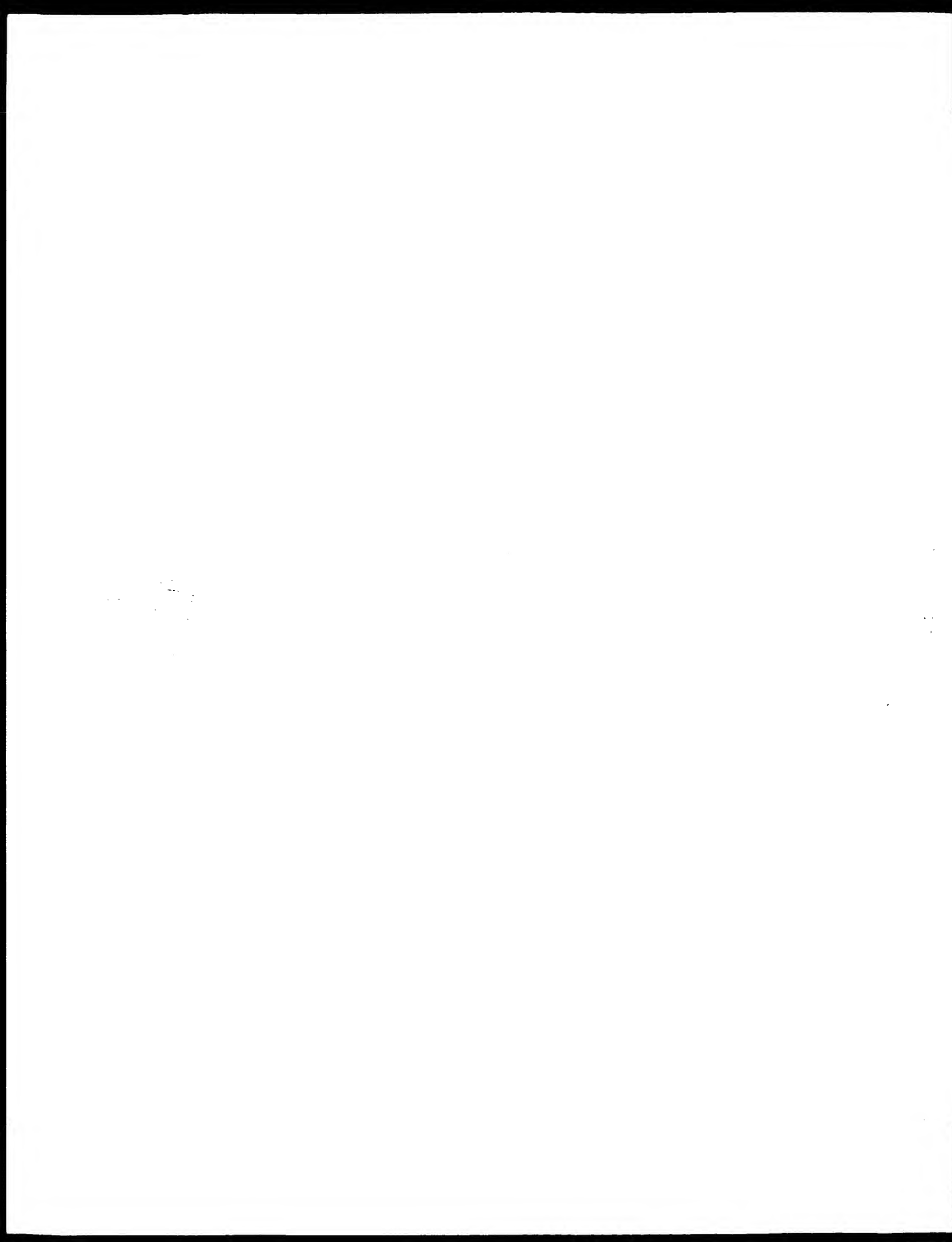












GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuLink Ltd.

AM nucleotide nucleotide search, using SW model

Run on: January 16, 2003, 15:20:22 : Search time 2966.48 seconds  
(without alignments)  
184669.506 Million cell updates/sec

Library: US-09-026-459A-48

Perfect scores: 4483

Sequence: 1 GCGGCAAGCGGGGAAAAA.....AAACAGATATATGACAGT 3383

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searches: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 420312

Minimum DB seq length: 0

Maximum Match: 90%

Maximum DB seq length: 200000000

Database: EST:\*

1: em\_estb1\*

2: em\_estb2\*

3: em\_estb3\*

4: em\_estb4\*

5: em\_estb5\*

6: em\_estb6\*

7: em\_estb7\*

8: em\_estb8\*

9: em\_estb9\*

10: em\_estb10\*

11: em\_estb11\*

12: em\_estb12\*

13: em\_estb13\*

14: em\_estb14\*

15: em\_estb15\*

16: em\_estb16\*

17: em\_estb17\*

18: em\_estb18\*

19: em\_estb19\*

20: em\_estb20\*

21: em\_estb21\*

22: em\_estb22\*

23: em\_estb23\*

24: em\_estb24\*

25: em\_estb25\*

26: em\_estb26\*

27: em\_estb27\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.2	26.9	1023	BM4500331	BM4500331 ACENCOURT
2	909.4	26.9	1051	BM561083	BM561083 ACENCOURT
3	896	25.3	1061	B0439221	B0439221 ACENCOURT
4	762.4	22.2	1122	BM543881	BM543881 ACENCOURT
5	725.4	21.4	727	AL597831	AL597831 DKFZP3130
6	699.6	20.7	824	B0519551	B0519551 DKFZP3130

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	692.2	20.5	767	12	B0529466
8	691.8	20.4	972	13	BM466194
9	678	20.0	845	14	BM454724
10	660	19.5	672	15	BM466859
11	643.6	19.0	797	16	BE539278
12	642.2	19.0	621	17	B0516219
13	638.8	18.9	1364	18	AK011246
14	617	18.2	618	19	BM264009
15	614	18.1	694	20	BE082846
16	596.4	17.6	922	21	B0529466
17	580	17.1	580	22	AL599105
18	575.2	17.0	914	23	AA763411
19	566.4	16.7	588	24	AA763411
20	564.9	16.7	881	25	B0529466
21	558	16.5	871	26	AA763485
22	530.2	15.7	598	27	AL582967
23	530	15.7	539	28	AL582967
24	528	15.6	572	29	AL582967
25	528	15.6	605	30	AL582967
26	524	15.5	623	31	AL582967
27	521.2	15.4	762	32	AL582967
28	518	15.3	832	33	AL582967
29	517	15.3	1154	34	BM802746
30	513.2	15.2	594	35	AV715544
31	512.4	15.1	709	36	AV715544
32	512.2	15.1	734	37	HE144857
33	505.2	14.9	736	38	BM151568
34	493.8	14.6	525	39	AA582967
35	493.4	14.6	597	40	BE931246
36	486.4	14.4	584	41	HE144857
37	484.4	14.3	494	42	HE144857
38	481.4	14.2	484	43	HE144857
39	480.2	14.2	659	44	BM151568
40	478.6	14.1	615	45	AA598846
41	474.4	14.0	734	46	B1529466
42	459.2	13.6	497	47	BM543881
43	457	13.5	457	48	AL597831
44	457	13.5	732	49	AL597831
45	455.4	13.5	643	50	B1554937

## ALIGNMENTS

RESULT 1  
BM4500331  
LOCUS: BM4500331, 1023 bp, mRNA, linear, EST 05 FEB 2002  
DEFINITION: ACENCOURT\_594544, NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5529637, 5% mRNA sequence.

ACCESSION: BM4500331, GI:18499071

VERSION: BM4500331

KEYWORDS: EST

SOURCE: Human

ORGANISM: Homo sapiens

REFERENCE: 1 (bases 1 to 1023)

AUTHORS: NIH-MGC http://mgc.nhlbi.nih.gov/

TITLE: National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL: Unpublished (1999)

COMMENT: Contact: Robert Strausberg, Ph.D.

Email: strausb@nhi.nih.gov

Tissue Procurement: ATCC/CRL-1073

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing: By: The J.M.A.G.E. Consortium (J.M.E.)

clone distribution: MGC clone distribution information can be

found through the J.M.A.G.E. Consortium/J.M.E. at:

Plate: LLAM1284 row: d column: 22

High quality sequence stop: 643

Location/Qualifiers: 1..1023

SOURCE: 1..1023

```

/clone_lib="NIH_MGC_72"
/tissue_type="mammary"
/lab_host="pH10B (phage-resistant)"
/clone="organ skin, Vector: pCMV-Sport6, Site_1: NotI,
Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size: 2 kb. Library constructed by Life
Technologies."
340 a 221 c 197 g 281 t

Query Match 26.9%; Score 910.2; DB 13; Length 1023;
Best Local Similarity 95.2%; Pred. No. 4, 1e-172;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

CY 1471 CAAGTGAAGATATATAGAGCTTAAATTCAAAATCATGTGTAAAGATATCTTCC 2030
DB 1 CAAGTGAAGATATATAGAGCTTAAATTCAAAATCATGTGTAAAGATATCTTCC 60
CY 2431 CAACTGCTTCACACACATTCACAACTGCTTTCATCAACACACACACATTCAC 2090
DB 1 CAACTGCTTCACACACATTCACAACTGCTTTCATCAACACACACACATTCAC 2090
CY 61 TCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 120
DB 1 TCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 120
CY 2491 TATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2150
DB 1 TATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2150
CY 121 TATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 180
DB 1 TATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 180
CY 2491 TATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2210
DB 1 TATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2210
CY 181 TTCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 240
DB 1 TTCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 240
CY 2211 TCCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 2270
DB 1 TCCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 2270
CY 241 TCCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 400
DB 1 TCCATGCTGCTCAGAGAAATTCACAACTGCTTTCATCAACACACACACATTCAC 400
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2330
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2330
CY 401 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 360
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 360
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2390
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2390
CY 601 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 420
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 420
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2450
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2450
CY 421 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 480
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 480
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2510
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2510
CY 481 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 540
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 540
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2570
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2570
CY 541 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 600
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 600
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2630
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2630
CY 601 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 660
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 660
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2690
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2690
CY 661 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 720
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 720
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2750
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2750
CY 721 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 780
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 780
CY 2491 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2808
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 2808
CY 781 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 840
DB 1 ATATAGTATTCATATAGCTGCTGCTTCATCAACACACACACACATTCAC 840

```

```

14 143 AATTTCAGAAATTTTGGAAAGAAATATTTTCAATATGCTTTTATGTCGCTGCTT 180
15 149 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 1368
16 181 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 240
17 169 TGGCTTTCGATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1428
18 243 TGGCTTTCGATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 300
19 1429 AATGAAAGGCTTTTAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1488
20 161 AATGAAAGGCTTTTAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
21 1489 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 1548
22 161 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 420
23 1449 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1548
24 1421 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 480
25 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1608
26 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 540
27 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 1728
28 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 600
29 1429 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1788
30 1401 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 660
31 1489 AAAAAAGGCTTTTAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1848
32 1461 AAAAAAGGCTTTTAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 720
33 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 1908
34 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 780
35 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1968
36 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 840
37 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2027
38 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 900
39 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
40 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
41 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
42 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
43 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
44 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
45 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
46 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
47 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
48 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
49 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
50 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
51 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
52 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
53 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
54 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
55 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
56 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
57 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
58 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
59 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
60 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
61 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
62 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
63 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
64 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
65 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
66 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
67 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
68 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
69 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
70 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
71 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
72 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
73 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
74 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
75 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
76 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
77 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
78 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
79 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
80 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
81 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
82 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
83 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
84 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
85 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
86 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
87 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
88 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
89 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
90 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
91 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
92 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
93 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
94 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
95 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
96 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953
97 1449 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2079
98 1421 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 953
99 1409 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2079
100 1481 GATATTAATGAAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 953

```

Tissue Procurement: Ailer  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The T.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Amersham Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the T.M.A.G.E. Consortium/LLNL at  
<http://image.llnl.gov>  
 Plate: LLNL4211 row: k column: 91  
 High quality sequence stop: 742.

BASE COUNT  
 347 a 190 c 184 g 371 t

Query Match 25.48; Score 856; Db 14; Length 1081;  
 Best Local Similarity 99.18; Pred. No. 40-161;  
 Matches 892; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

QY 2485 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 2644  
 Db 1 GAGGCTGTAATGGGCAATATAAGCAAGCAATACATTCAGATTCGACACAT 60

QY 2545 ATGACTTCTACGGAACAGCAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2604  
 Db 61 ATGACTTCTACGGAACAGCAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 120

QY 2605 AATGAAAGGCAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2664  
 Db 121 AATGAAAGGCAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 180

QY 2665 CATTGCTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2724  
 Db 181 CATTGCTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 240

QY 2725 ATATCTTACGCTCTTTTGGGCAATATAAGCAAGCAAGCAAGCAAGCAAGCAAG 2784  
 Db 241 ATATCTTACGCTCTTTTGGGCAATATAAGCAAGCAAGCAAGCAAGCAAG 400

QY 2785 AATGAAAGGCAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2844  
 Db 401 AATGAAAGGCAATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 460

QY 2845 TGGCATTTAAAAAGTTGTAAGCAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2904  
 Db 361 TGGCATTTAAAAAGTTGTAAGCAAGTTGTTGTTGTTGTTGTTGTTGTTGTT 420

QY 2905 TGGCATTTAAAAAGTTGTAAGCAAGTTGTTGTTGTTGTTGTTGTTGTTGTT 2964  
 Db 421 TGGCATTTAAAAAGTTGTAAGCAAGTTGTTGTTGTTGTTGTTGTTGTTGTT 480

QY 2965 TGGCATTTAAAAAGTTGTAAGCAAGTTGTTGTTGTTGTTGTTGTTGTTGTT 3024  
 Db 481 TGGCATTTAAAAAGTTGTAAGCAAGTTGTTGTTGTTGTTGTTGTTGTTGTT 540

QY 3025 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3084  
 Db 541 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600

QY 3085 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3144  
 Db 601 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660

RESULT 4  
 QY 49221  
 LOCUS A9221\_276556.2 NIH MGC 92 Homo sapiens cDNA clone IMAGE:6015480  
 VERSION QY 49221  
 KEYWORDS 531  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catartini; Hominidae; Homo;  
 1 (bases 1 to 1081)  
 NIH MGC <http://image.llnl.gov/>  
 National Institutes of Health  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [caples@mail.nih.gov](mailto:caples@mail.nih.gov)





## COMMENT

Contact: Poustka A.J.  
 Department: Leibniz  
 Max Planck Institute for Molecular Genetics  
 Ihnestrasse 73, 14195 Berlin, Germany  
 Tel.: +49-30 84141624  
 Fax: +49-30 84141129  
 Email: poustka@mpi-berlin-dahlem.mpg.de  
 This is the 5' sequence of the clone insert.  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by DKFZ (German Cancer Research Center,  
 Heidelberg/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 No SI sequence available.  
 This clone (DKFZp3160816) is available at the RZPB in Berlin.  
 Please contact the RZPB: Ressourcenzentrum, Houbnerweg 6, 14059  
 Berlin Charlottenburg, GERMANY; Email: clone@rzpb.de.

## FEATURES

LOCUS 1..727  
 DEFINITION /organism="Homo sapiens"  
 ACCESSION /db\_xref="taxon:9606"  
 VERSION /clone="DKFZp3160816"  
 KEYWORDS /clone\_lib="413 (synonym: hicc2)"  
 SOURCE /dev\_stage="adult"  
 ORGANISM /lab\_host="DH10B"  
 REFERENCE /note="Vector: pRipEx2; Site\_1: SIIIA; Site\_2: SIIIB;  
 AUTHORS cDNA-collection"  
 JOURNAL  
 COMMENT  
 BASE COUNT 248 a 163 c 130 g 198 t  
 ORIGIN  
 Query Match 21.4% Score 7.25.4; DB 9; Length 727,  
 Best Local Similarity 99.9% Pred. No. 4.4e-195;  
 Matches 72% Conservative 9; Mismatches 1; Indels 0; Gaps 9;  
 QY 1761 GAAATACCTGCTGCTTACACCTTTTATAAAAAAGCTGATGGGTAGCTATCCCGCT 1820  
 Db 1 GAAATACCTGCTGCTTACACCTTTTATAAAAAAGCTGATGGGTAGCTATCCCGCT 60  
 QY 1821 AAAATACATTTTGCAATGCTTTCGCTGAGCAGCCAGCAATACAGATATCTGGAC 1880  
 Db 61 AAAATACATTTTGCAATGCTTTCGCTGAGCAGCCAGCAATACAGATATCTGGAC 120  
 QY 1881 GTTTTCAGACATGCTGAGATGATGATGAAATCATGAGACAGACATTTGGACCA 1940  
 Db 121 GTTTTCAGACATGCTGAGATGATGATGAAATCATGAGACAGACATTTGGACCA 180  
 QY 1941 AATTAAGCATGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 2000  
 Db 181 AATTAAGCATGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 240  
 QY 2001 AATTAAGCATGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 2060  
 Db 241 AATTAAGCATGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 300  
 QY 2061 TTTGATGAAAAGAGAGATGATGATTTATATATGATTTTATATGATGCTGCTTCATGCA 2120  
 Db 601 TTTGATGAAAAGAGAGATGATGATTTATATATGATTTTATATGATGCTGCTTCATGCA 360  
 QY 2121 GAGATTTAAAAAATAATTTTGGATATGCTTCCAGCAGGCGCGTACCTTTGTCACCAAT 2180  
 Db 361 GAGATTTAAAAAATAATTTTGGATATGCTTCCAGCAGGCGCGTACCTTTGTCACCAAT 420  
 QY 2181 AGCTGATATGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCTG 2240  
 Db 421 AGCTGATATGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCTG 480  
 QY 2241 GAAATATATATTTTAAAGATGAAATGATATATATATATATATATATATATATATATAT 2300  
 Db 481 GAAATATATATTTTAAAGATGAAATGATATATATATATATATATATATATATATATAT 540  
 QY 2301 AATAAAAATGATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCTG 2360  
 Db 541 AATAAAAATGATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCTG 600

QY 2361 TGAGAAATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2420  
 Db 601 TGAGAAATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660  
 QY 2421 TGCTGAAGAAAGCAAGCTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2480  
 Db 661 TGCTGAAGAAAGCAAGCTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720  
 QY 2481 ACATGAA 2487  
 Db 721 AGATGAA 727  
 RESULT 6  
 LOCUS HG610661  
 DEFINITION 602611848T1 NIH\_MGC\_66 Homo sapiens cDNA clone IMAGE:4747634 5',  
 mRNA sequence.  
 ACCESSION HG610661  
 VERSION HG610661.1 GI:14662032  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 823)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rcp@bbs-rmail.nih.gov  
 Issue Procurement: NCI/NIH  
 cDNA library preparation by: CLONTECH Laboratories, Inc.  
 DNA library Arrayed by: the I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 plate: L1CM1605 row: 0 column: 11  
 High quality sequence stop: 709.  
 Location/Qualifiers  
 1..823  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4747634"  
 /clone\_lib="NIH\_MGC\_60"  
 /issue\_type="adenocarcinoma"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: prostate; Vector: pDRK-lib (Clontech);  
 Site\_1: SfiI (480-500bp); Site\_2: SfiI (480-500bp);  
 ); double-stranded cDNA was prepared from cell line RNA,  
 5' and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CATGGCATATATGCT-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGGCGAGGCGGCAATG dT(0)EN-3'  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
 Library."

## BASE COUNT

262 a 148 c 142 g 271 t

## ORIGIN

Query Match 20.7% Score 699.6; DB 12; Length 824;  
 Best Local Similarity 94.8% Pred. No. 6.5e-130;  
 Matches 779; Conservative 0; Mismatches 34; Indels 9; Gaps 5;  
 QY 2511 AGCAGACTCCAAATTTTCACGACAAACCTGGCAGCAAAATGATTTCTACGACAAATGAAATGAA 2570  
 Db 1 AGCAGACTCCAAATTTTCACGACAAACCTGGCAGCAAAATGATTTCTACGACAAACGAAATGAA 60  
 QY 2571 AAAGCAATAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2630







```

27 2848 CATTAAAGAAATTTGTAAGAGATTCTTCTTGGCAAGAGAAAGAAATATGCTGTTATAG 2907
DB 601 CATTAAAGAAATTTGTAAGAGATTCTTCTTGGCAAGAGAAAGAAATATGCTGTTATAG 660
27 2908 AATGTAAGAAAG 2919
DB 601 AATGTAAGAAAG 672

RESULT 11
LOCUS BE549278 797 bp mRNA linear EST 09-AUG-2000
DEFINITION BE549278.1 NIH MGC_51 Homo sapiens cDNA clone IMAGE:3449873.5;
mRNA sequence.
ACCESSION BE549278
VERSION BE549278.1 GI:39767924
KEYWORDS EST.
SOURCE Human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 797)
AUTHORS NIH MGC http://mgi.nhl.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstrauss@mail.nih.gov
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the L.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8427 row: 1 column: 18
High quality sequence stop: 696.
FEATURES
Source 1..797
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3449873"
/clone_lib="NIH MGC_10"
/contig="MGC_51"
/lab_host="EBLOR"
/notes="Organ: cervix; Vector: pCMV Sp6R6; Site: 1; Not1;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 242 a 171 c 167 g 207 t
ORIGIN
Query Match 19.0%; Score 643.6; DB 10; Length 797;
Best local similarity 95.8%; Prod. No. 11c-118;
Matches 747; Conservative 0; Mismatches 24; Indels 9; Gaps 8;
27 2046 CTGTTAGAGAGATTCAAAAGTCTTTGATCAAGAGAGAGAGATGATCTATTATAG 2095
DB 1 CTGTTAGAGAGATTCAAAAGTCTTTGATCAAGAGAGAGAGATGATCTATTATAG 59
27 2096 TATGCTAATATGAGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2155
DB 60 TATGCTAATATGAGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 118
27 2156 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2215
DB 119 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 178
27 2216 GTTACAGGCTTACGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2275
DB 179 GTTACAGGCTTACGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 238
27 2276 AATTTAGAGAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2335
DB 111 AATTTAGAGAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT

```

```

DB 239 AATTTAGAGAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 298
QY 2396 CAATTGGCAATCATTCGGACATCTTCAGAAATTCAGAAAAAATAAATACAGAGAGAGAGAG 2495
DB 299 CAATTGGCAATCATTCGGACATCTTCAGAAATTCAGAAAAAATAAATACAGAGAGAGAGAG 458
QY 2396 ACAGAGAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2454
DB 359 ACAGAGAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 418
QY 2455 AATCTAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2514
DB 419 AATCTAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 478
QY 2515 GAGTCTCAAAATTCAGAGCAAAATTCAGAGCAAAATTCAGAGCAAAATTCAGAGCAAAATTC 2574
DB 479 GAGTCTCAAAATTCAGAGCAAAATTCAGAGCAAAATTCAGAGCAAAATTCAGAGCAAAATTC 538
QY 2574 GCAGAAATTAATGATAGTATGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2634
DB 539 GCAGAAATTAATGATAGTATGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 598
QY 2634 CTTCTGGTGGACATCTGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2694
DB 599 CTTCTGGTGGACATCTGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 656
QY 2694 CTTCTGGTGGACATCTGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2754
DB 657 CTTCTGGTGGACATCTGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 714
QY 2754 ATGTTGAGATGATATGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2814
DB 715 ATGTTGAGATGATATGATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 774
RESULT 12
LOCUS BG616219 681 bp mRNA linear EST 18 APR 2001
DEFINITION BG616219.1 NIH MGC_51 Homo sapiens cDNA clone IMAGE:4775342.5;
mRNA sequence.
ACCESSION BG616219
VERSION BG616219.1 GI:134667590
KEYWORDS EST.
SOURCE Human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS NIH MGC http://mgi.nhl.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstrauss@mail.nih.gov
cDNA Library Preparation: ATCC
cDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the L.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1650 row: 6 column: 65
High quality sequence stop: 679.
FEATURES
Source 1..681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4775342"
/clone_lib="NIH MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="JHI10B (11 plaque resistant)"
/notes="Organ: Testis; Vector: pCMV LTR (Clontech); Site: 1;
SL12 (qacc:ctcacc); Site 2: SL11 (qacc:atgacc);
Double-stranded cDNA was prepared from cell line RNA.

```









from GRESTES PCR (U.S. Letters Patent application No. 196  
 2716 - Ludwig Institute for Cancer Research) profiles  
 into the p9C18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions.

BASE COUNT 411 G 129 A 127 T 226 C 1 others

Query Match 18.16; Score 614; DB 10; Length 694;

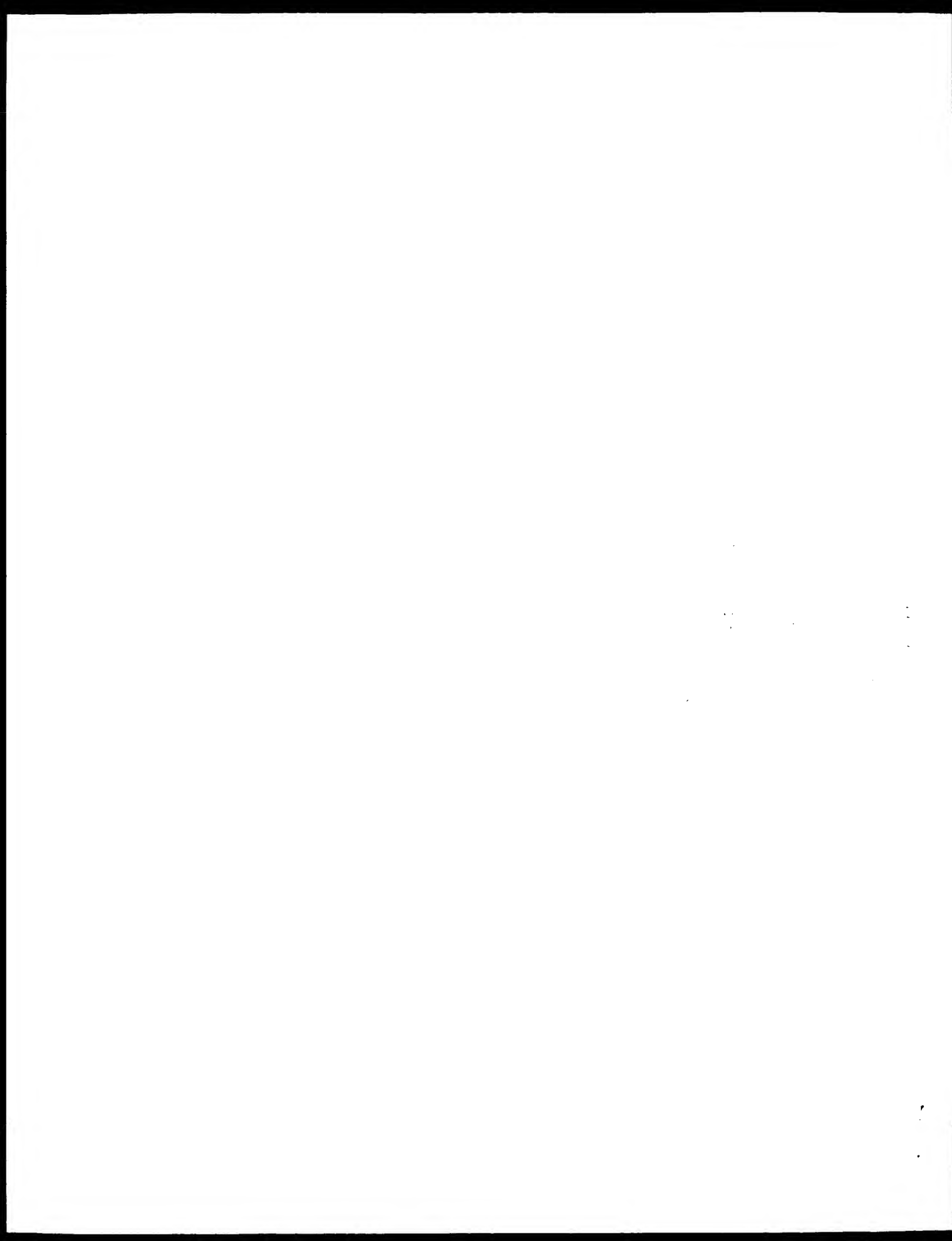
Best Local Similarity 97.48; Prod. No. 9.3e-113;

Matches 64%; Conservative 9; Mismatches 16; Indels 2; Gaps 2;

```

27 1044 TTATTACAACTGCAATGTAATGCAAAAGAAAGTATATCTGAAAGAGGTGAAGGATATAGG 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 683 TATTTAAGACATGACAGC AICCAAGAAAGATATATGAGAGAGCAAGCAATAGG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1104 ATATATCTTTAAAGAGAAATTTGCTAAAGCTGGGAGACAGGTGGTCAAAATGGCAG 1163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 624 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1164 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 504 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1224 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 504 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1284 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 444 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1344 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 484 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1404 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 424 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1464 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 264 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1524 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 204 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1584 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 144 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1644 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 1702
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 84 ATATATCTTTAAAGAGAAATTTGCTAGAGCTGGGAGACAGGTGGTCAAAATGGCAG 25
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 1704 AT 1705
    |||
60 24 AT 22
  
```

Search completed: January 16, 2003, 04:11:17  
 Job Time: 29.09.56.8008





```

AR091965      AR091965      2994 bp      DNA      linear      PAT 08-SEP-2000
LOCUS       Sequence 4 from patent US 5998134.
DEFINITION  AR091965
VERSION     AR091965.1 GI:10018719
KEYWORDS   Unknown.
SOURCE     Unpublished.
ORGANISM   Homo sapiens.
REFERENCE   1 (bases 1 to 2994)
AUTHORS   Lee, W.-H. and Lee, F.-Y.-H. P.
TITLE     Retinoblastoma gene-cancer suppressor and regulator
JOURNAL   Patent US 5998134-A 4 07-DEC-1994;
FEATURES             Location/Qualifiers
     source          1..2994
                     /organism="unknown"
     BASE COUNT      974 a 618 c 594 g 808 t
     ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
```



Db	799	GGGCGCTTCAGCTATTTTATTAATTAACCTCTCACTTCCCATGTTGGCTCAAGAACCATATAAA	858
QY	241	Thr-----GlySer-----	243
Db	859	ACAGCGCTTATACCGCAATAAAGGCTCACTTCAGACACCCAGCCAGGCTCAGAACACAGCACT	918
QY	243	-----	243
Db	919	GTATGATATATCAAAAATATATAAAAATATATCAAGAAATATATGAAGTCTCTCTGTAAAGAA	978
QY	243	-----	243
Db	979	CAATCAATGCTAAATATACATGAGCTGAAAAATGTTTATTTCAAAAATTTTATACCTTTTAAG	1038
QY	244	AspSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGluLeuLeuSerTyrAsn	263
Db	1039	AAATCTCTTCAGCTTCGTAAACATCAATAGCGACTTCAGAGGTGTGAAAACTCTCTTCAACGA	1098
QY	264	TyrGluGluLeuLeuTyrLeuLysAsnLysAspLeuAspAlaIleArgLeuPheLeuAspHisAsp	283
Db	1099	TACGAAATAATTTATCTTAAAAATAAATCATCATGTCAGATATATTTTGGATCATGAT	1158
QY	284	LysThrLeuGluThrAspSerIleAspSerPheGluThrGlnArgThrProArgLysSer	303
Db	1159	AAAACTCTTCAGACTCATCTCATACACAGTTTGTGAACACACAGACACACCACGAAAAAGT	1218
QY	304	AsnLeuAspGluValAsnValIleProProHisThrProValAlaArgThrValMetAsn	323
Db	1219	AAGCTTCATGACACAGCTGAATGTAATCTCTCCACACATCCAGTTAGCGATGTATTGAAC	1278
QY	324	ThrIleGluGluLeuMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu	343
Db	1279	ACTATGCAACAAATTAATGATATTTTAAATTCAGTAATGATAAATCTTCAGAAATCTG	1338
QY	344	IleSerTyrPheAsnAsnCysThrValAsnTrolLysGluSerIleLeuLysArgValLys	363
Db	1339	ATTTCCATTTTAAACACTCCACACTGAATCCAAAGAAAGATATACTGAAAACAGCTCAAG	1398
QY	364	AspIleGlyTyrIlePheLysGlnLysPheAlaLysAlaValGlyGluGlyCysValGlu	383
Db	1399	GAATAGATATATCTTTAAACAGAAATTTGCTAAAGCTGTGGCAGAGGTTGTGTCCAA	1458
QY	384	IleGlySerGluArgTyrLysLeuGlyValArgLeuTyrTyrArgValMetGluSerMet	403
Db	1459	ATTGGATCACAGCATATANAATTAGATTGTTGTTATTATTAATTAATTAATTAATTAATG	1518
QY	404	LeuLysSerGluGluArgLeuSerIleGlnAsnPheSerLysLeuLeuAsnAspAsn	423
Db	1519	CTTAAALCACAGCAAGACCATTCACATCAAAATTTTACCAAACTTCGAAATGACAC	1578
QY	424	IlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrTyrSerArg	443
Db	1579	ATTTTCATATGCTTTATGCGGCTGCGCTTCAGCTGTGTAATGGCCACATATAGCAGA	1638
QY	444	SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProThrPheLeuAsnVal	463
Db	1639	AGTACATCTCAAACTTATTTTAAATAGATTGTCTTTTCGATTCGATTCGATGTC	1698
QY	464	LeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPheIleLysAlaGluGly	483
Db	1699	CTTAAATTAAGACCTTCATCTTTTACAACTGATCCGAAAGCTTTATCAAGCAGAGGCG	1758
QY	484	AsnLeuThrArgGluMetIleTyrLysIleLeuGluArgCysGluHisArgTleMetGluSer	503
Db	1759	AACATGACAGCAAAATGATAAACATTTTGAACGCATGTCAACATCCGATTCGAAATCC	1818
QY	504	LeuAlaIlePheSerAspSerPheLeuPheAspLeuIleLysGluSerLysAspArgGlu	523
Db	1819	CTTGATGGCTCTCAGATTCACCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT	1878
QY	524	GlyProThrAspHisMetGlnSerAlaCysPheLeuAsnLeuProGluGlnAsnAsnHis	543

DQ	1879	GGAACGAACTGATCATTTTAAAGCCTGGTTCGTAATCTCCGCGAATAAATAC	1938
OY	544	ThrAlaAlaAspMet.TyrLeuSerProValAArgSerProLysLysLysGlySerThrThr	563
DB	1939	ACTGTAGCAATATGATCTTTCTCTGTAGATCTCCAAAGAANAAGGTTTAACTACG	1998
OY	564	ArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLys	583
DB	1999	CGTGTAATATCTATCTCAAATGAGAGAAAGAAATGAGATCTGAGATCTGAGATCTG	2058
OY	584	ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysVal.TyrArgLeuAlaTyrLeu	603
DB	2059	CCATTGAAATCTACCTCTCTTTCACATGTTTTATAAAAAAGTGATCGCTAGCCTATCTC	2118
OY	604	ArgLeuAsnThrLeucylsGluArgLeuLeuSerGlnHisProGluLeuGlnHisIleLe	623
DB	2119	CGGTAAATATCACTTTGTCAAGGATCTGTGTATGAAATAGAACATATCATCTC	2178
OY	624	TripThrLeuPheGlnHisThrLeuGlnAsnGluTyrGluLeuMetAlaGspArgHisLeu	643
DB	2179	TGACACCTTTTTCAGCACACCTCTGAGATCATGATGTAATCTATGAGTAGAGATTTG	2238
OY	644	AspGlnIleMetMetCysSerMetTyrGlyIleCysLysValLysAsnIleAspLeuLys	663
DB	2239	GACTAAATATACATGCTGCTGCATGATGATGATGATGATGATGATGATGATGATG	2298
OY	664	PheLysIleIleValThrAlaTyrLysAspLeuProHisAlaValGlnGluThrPheLys	683
DB	2299	ITCAAATCAITTTATACAGCAIAAAGATCTTTCTGATGATGATGATGATGATGATG	2358
OY	684	ArgValIleLysGluGluGluTyrAspSerIleIleValPheTyrAsnSerValPhe	703
DB	2359	CGTGTTTATATCAAAAAGAGATATGATTTATATATGATTTCTATATCTGGCTCTC	2418
OY	704	MetGlnArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArgProProThrLeuSer	723
DB	2419	ATTCAGACATCAAAACAATAATTTTCTGATATCTCTCCACCAGCCCCCACTCTCCA	2478
OY	724	ProIleProHisIleProArgSerProTyrLysPheProSerSerProLeuAlaIlePro	743
DB	2479	CATAATCTTAATCTGAGAGGCCCTTAAAGTTTCTAGTTTCTAGTTTCTAGTTTCT	2538
OY	744	GlyCylAsnIleTyrIleSerProLysLysSerProTyrLysIleSerGluCylLeuPro	763
DB	2539	GGAGGGAAACATATATTTTCACTCTGAAATATATATAAATTTAGAGTCTCTCTCA	2598
OY	764	ThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlyGluSerPheLys	783
DB	2599	ACACCACAAAAAGATCTCAACATCAACAATCTAGTATCAATGCTGTAATCAATCTG	2658
OY	784	ThrSerGluLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArgValLeuLys	803
DB	2659	AATTCTGAAAGTTTCAAAAATAATATAGATGGATATGATAGAGATCTGAGTATCTG	2718
OY	804	ArgSerAlaGluCylSerAsnProProLysProLeuTyrLysLeuArgPheAspIleGlu	823
DB	2719	AGAAGTCTGAGAGAGACAAATCTCTAAATCAATGAAAAACTATGTTTATATAGAA	2778
OY	824	GlySerAspGlnAlaAspLysSerIleHisLeuProGlyLysPheGlnGlnLys	843
DB	2779	GGATCGATCAAGACAGATCGAATCAACATCTCCACAGAGATCTCAATTTTCACAGAAA	2838
OY	844	LeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetAsnAspSerMet	863
DB	2839	CTGTAATAATATATTTACTCTCAATCAATATCAATCAATCAATCAATCAATCAATCA	2898
OY	864	AspThrSerAsnLysGluGlnLys 871	
DB	2899	CATACCTCAAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	2923











344 AsnLeuThrArgIleMetIleLysHisLeuAlaArgCysGlnHisArgIleMetIleGluSer 503  
 1769 AATCTGCAACACAAATATATAAAATATATACCAAGATCGCAATCGCAATCGCAATCG 1818  
 504 LeuAlaLeuProSerAspSerProLeuLeuAspLeuIleLysGlnSerLysAspArgGlu 523  
 1819 CTCTATGATCTCTAGAT 1878  
 524 GlyProThrAspHisLeuSerAlaCysProLeuLeuAspLeuIleLysGlnSerLys 543  
 1879 GAAACAAAT 1938  
 544 ThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLysLysLysLysLys 563  
 1939 AATGACAT 1998  
 564 ArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLys 583  
 1999 GATGTAAT 2058  
 584 ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyrArgLeuAlaTyrLeu 603  
 2059 GATGTAAT 2118  
 604 ArgLeuAsnThrLeuCysGlnAlaLeuLeuSerGlnHisProGlnLeuGlnHisIleIle 623  
 2119 GATGTAAT 2178  
 624 TrpThrLeuPheHisThrLeuGlnAsnGluTyrGlnLeuMetArgAspArgHisLeu 643  
 2179 TATGTAAT 2238  
 644 AspGlnIleMetMetCysSerMetTyrGlyIleCysLysValLysAsnIleAspLeuLys 663  
 2239 GATGTAAT 2298  
 664 PheLysIleIleValThrAlaTyrLysAspLeuProHisAlaValGlnGlnThrPheLys 683  
 2299 TATGTAAT 2358  
 684 ArgValLeuIleLysLysGlnGluIleTyrAspSerIleIleValPheTyrAsnSerValPhe 703  
 2359 GATGTAAT 2418  
 704 MetGlnArgLeuLysThrAsnIleGlnGlnIleTyrAlaSerThrArgProArgHisLeuSer 723  
 2419 AATGTAAT 2478  
 724 ProLeuProHisIleProArgSerProTyrLysPheProSerSerProLysArgIlePro 743  
 2479 GATGTAAT 2538  
 744 GlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysLysSerGlnIleLeuPro 763  
 2539 GATGTAAT 2598  
 764 ThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlnGlnSerPheGly 783  
 2599 AATGTAAT 2658  
 784 ThrSerGlnLysPheHisThrIleAsnGlnMetValLysAsnSerAspArgValLeuLys 803  
 2659 AATGTAAT 2718  
 804 ArgSerAlaIleGlnLysSerAsnProProLysProLysLysLysLysLysLysLys 823  
 2719 AATGTAAT 2778  
 824 GlySerAspGlnAlaAspGlnLysLysLysLysLysLysLysLysLysLysLysLys 843  
 2779 GATGTAAT 2838

844 LeuAlaGluMetThrSerThrArgIleMetIleLysGlnHisArgIleMetIleGluSer 863  
 2839 CTCTATGATCTCTAGAT 2898  
 864 AspThrSerAsnLysGlnGlnLys 871  
 2899 GATGTAAT 2922  
 RESULT 6  
 HUMRBSA 2994 bp mRNA linear 04-12-JUL-1996  
 LOCUS Human retinoblastoma susceptibility protein mRNA, complete cds.  
 DEFINITION M28419.1 GI:190962  
 ACCESSION M28419  
 VERSION 1 (bases 1 to 2994)  
 KEYWORDS retinoblastoma protein; retinoblastoma susceptibility protein.  
 SOURCE Homo sapiens retinoblastoma cDNA to mRNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
 1 (bases 1 to 2994)  
 REFERENCE 1  
 AUTHORS Lee, W.H., Shew, J.Y., Hong, F.D., Sery, L.W., Denos, L.A., Young, L.J.,  
 Bookstein, R., and Lee, E.Y.  
 TITLE The retinoblastoma susceptibility gene encodes a nuclear  
 phosphoprotein associated with DNA binding activity  
 JOURNAL Nature 329 (6140): 642-645 (1987)  
 MEDLINE 88014248  
 PubMed 3657987  
 FEATURES  
 location/Qualifiers  
 1..2994  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="Xp22.42:Yp11.3"  
 /cell\_line="retinoblastoma cell line Y79, and normal  
 retinoblastoma cell line"  
 /tissue\_type="retinoblastoma"  
 1..2994  
 /gene="M1C2"  
 149..12925  
 /gene="M1C2"  
 /note="retinoblastoma susceptibility protein"  
 /codon\_start 1  
 /protein\_id="AA069808.1"  
 /db\_xref="GI:190962"  
 /db\_xref="Gene:600 120-184"  
 /translation="MEPKIPKIAAIAAAAAAEPAPPPPPPPPPPPPPPPPPPPPPPPPPPPPP  
 VRIETLLEPLFALCAKLTGVRKAWLIMKAVSSVIVLALYKKKLEWGLP  
 IFIAADIDEMSFTELOKNIISVHKFFNLKEDISIKVHNSRLKKKYDLEA  
 LESKLETCELYLTQPSSTSTEISALVLSWLIILAKGEVLQMEDQVLSOL  
 MCVLDYFIKSPMLKEPKYTAVIPINGSPTRPKQNSARLAKULENDRLIEV  
 LKKEICNDIEVKNVYFNFTFFMNSLGLVTSNGLPEVENLSKRTFEIYLNKRLDAR  
 LKDHOKLTQDSTDSFQTPRKSNIDERNVITPPHPVPTVNTITQOIMMLNSA  
 SPQSNLSYFNCTVNPKNLSLKKKIDGYIFKRFKAVGQGVVIGTSQRYKGV  
 RLYRYMSEMKSEERLSTONSKLLNIFHMSLALAEVMAIVSRISLONLIS  
 GLDSEFWLNVNLKALFYKVESEFKAFENLTREMLKHLERLHRIMESLAWLS  
 SPLEDLTKQKURESPDHLFACPLNPLONNHITAAIMYISVPSPKKSGSTIRVNS  
 TANAEFGATSAETQKPKSTSLSEFYKVKYKLVLRLNTILEMLSEIPELEHITWT  
 LKQHLQNSFYFMRHQTQIMMSYGLCKVKNLIDKFFLIVAYKDLHVAQVPEK  
 RVTKKEFYDSITVEYNSVMORKNITQASRPDLSLPHIPSPDPKPSPLER  
 DPGNLYTSPLASPKLSGLPLPKMLPRKPLTSGESFQSTSRFQKINOMVNSD  
 RVLKRSAGSNPKPKLKEFDIERDEARGLKLPKESKFOOKLAEMTSTRTMQU  
 KNSDSNLSNKEEK"  
 BASE COUNT 974 a 618 c 593 q 809 t  
 ORIGIN  
 Alignment Scores:  
 Fred. No.: 2.94e-410 Length: 2994  
 Score: 4447.50 Matches: 871  
 Percent Similarity: 93.86% Conservation: 0  
 Best Local Similarity: 93.86% Mismatches: 0  
 Query Match: 98.92% Indels: 57  
 DH: 9 Gaps: 2  
 US-09-026-459A-49 (1-871) x HUMRBSA (1-2994)





QY 181 IleSerThrGluIleAsnSerAlaLeuValLeuIysValSerTrpIleThrPheLeuLeu 200  
 DB 544 ATAATCTACTCAAAATAAATTTGGATTGGTGTAAAAAGTTCTTGGATCACATTTTATTA 603  
 QY 201 ALILysGlyGluValLeuGlnMetGluAspLeuValIleSerPheGlnLeuMetLeu 220  
 DB 604 GCATAAGGCAAGATTTACAAATTCGAAGATGATCTGCTGATTTTCATTCAGTTAATGCTTA 663  
 QY 221 CysValIleAspThrPheIleLysLeuSerProProMetLeuLeuLysGlnProTrpTyrLys 240  
 DB 664 IRTGATCTTAAATATTTTATTAATATCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723  
 QY 241 Thr-----CysSer----- 243  
 DB 724 ACACATTTTATATGATTAATGCTTTCATCTGAACTACCTGAGGCTGCTGCTGCTGCTGCTGCT 783  
 QY 243 ----- 243  
 DB 784 GCACGGCAACCAAAACACATGAGAAATGATACCAAGATTTATGCAAGTTCTCTCTCTAAAGAA 843  
 QY 243 ----- 243  
 DB 844 CATCAATCTAATATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903  
 QY 244 AsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValIleAsnLeuSerLysArg 263  
 DB 904 AATTCCTTGGATTTCTATATCTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963  
 QY 264 TyrIleGlnIleTyrLeuLysAsnLysAspLeuAspAlaArgLeuPheLeuAspHisAsp 283  
 DB 964 IACCAAGAAATTTATCTTTAAAAATAAAGATCTAGATGCAAGATATTTTTCGATCATGAT 1023  
 QY 284 LysThrLeuGlnThrAspSerIleAspSerPheGluThrGlnArgThrProArgLysSer 303  
 DB 1024 AAAATCTTCAGACATGATCTATAGACATTTTGAACACACAGAACACCCAGCAAAAAAGT 1083  
 QY 404 AsnLeuAspGluGluValAsnValIleProProHisIleThrProValArgThrValMetAsn 323  
 DB 1084 AAGCTTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143  
 QY 424 ThrIleGlnGlnLeuMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 343  
 DB 1144 ACATCAACAAATTAAGATGATTTTAAATTCAGCAAGTATGATCAACCTTCAGAAAATCTG 1203  
 QY 444 IleSerTyrPheAsnAsnCysThrValAsnProLysGluSerIleLeuLysArgValLys 363  
 DB 1204 ATTCTGATTTTAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1263  
 QY 464 AspIleGlyTyrIlePheLysGluLysPheAlaLysAlaValIleGlnCysValGlu 383  
 DB 1264 GATATAGATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1323  
 QY 294 IleLysSerGlnArgTyrLysLeuGlyValArgLeuTyrArgValMetGluSerMet 403  
 DB 1324 ATGGATCAACAGCAATCAAACTTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1383  
 QY 404 LeuLysSerGluGluArgLeuSerIleGlnAsnPheSerLysLeuLeuAsnAspAsn 423  
 DB 1484 CTAAATACAGAAAGCAAGGATATGATCTCAAAATTTTGAAGAAATCTCTGAAATGACAAAC 1443  
 QY 424 PhePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrTyrSerArg 443  
 DB 1444 ATTTTCTATATGCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1503  
 QY 444 SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProTrpIleLeuAsnVal 463  
 DB 1504 AGTACATCTCAGAAATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1563  
 QY 464 LeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPheIleLysAlaGluLys 483  
 DB 1564 CTAAATTTAAAGCCCTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1623  
 QY 484 AsnLeuThrArgGluMetIleLysHisIleGluArgCysGlnHisArgIleMetGluSer 503

DB 1624 AACTTCACAAAGACAAATGATATAAAATTTTAGAACGATGUGAACATCGAAATCAATCAATCC 1683  
 QY 504 LeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAsnArgGln 523  
 DB 1684 CTGTGATGCTCTCAGATTCACCTTTTATTTGATCTTATTAACAAATCAAAAGCAGCGAGAA 1743  
 QY 524 GlyProThrAspHisLeuGluSerAlaCysPheLeuAsnIleProGluGlnAsnAsnHis 543  
 DB 1744 GAGCAAAATTCATATTTTAAATTTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1803  
 QY 544 ThrAlaAlaAspMetTyrLeuSerIleValArgSerProLysLysLysLysLysLysLysLysLys 563  
 DB 1804 ACTCCATCATATATGATCTTCT 1863  
 QY 564 ArgValAsnSerThrAlaAsnAlaGlnThrGlnAlaThrSerAlaPheGlnThrGlnLys 583  
 DB 1864 CGTGAATTTCTACTGCTCAAAATGACAGACACAAAGCAAACTAGCTTCTCTCTCTCTCTCTCTCT 1923  
 QY 584 ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyrArgLeuAlaTyrLeu 603  
 DB 1924 CATTAAATTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1983  
 QY 604 ArgLeuAsnThrLeuGlnCysGluArgLeuLeuSerGlnHisProGluLeuGlnHisIleIle 623  
 DB 1984 CGCTAAATACACTTTCTGACGCT 2043  
 QY 624 TrpThrLeuPheGlnHisThrLeuGlnAsnGluTyrGluLeuMetArgAspArgHisLeu 643  
 DB 2044 TGCACCTTTCT 2103  
 QY 644 AspGlnIleMetMetCysSerMetTyrGlyIleCysLysValLysAsnIleAspLeuLys 663  
 DB 2104 GACCAAAATATGATGTTTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2163  
 QY 664 PheLysIleIleValThrAlaTyrLysAspLeuProHisAlaValGlnGlnThrPheLys 683  
 DB 2164 TTTCAAAATCATTTGTAACAGCATACAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2223  
 QY 684 ArgValLeuIleLysGluGluTyrAspSerIleIleValPheTyrAsnSerValPhe 703  
 DB 2224 CGTCTTTCT 2283  
 QY 704 MetGlnArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArgProThrLeuSer 723  
 DB 2284 ATGCAGAGATCTGAAACAAATATTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2343  
 QY 724 ProIleProHisIleThrArgSerTrpTyrLysIleProSerSerProArgIlePro 743  
 DB 2344 CCAATACCTTCACATCTCTCTGAGCCCTTACAAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2403  
 QY 744 GlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysIleSerGlnGlyLeuPro 763  
 DB 2404 GGAGGGAACATCTATATTTTCCCTCTGAGAGCTCCATATAAAAAATTCACAAAGCTCTCTCTCT 2463  
 QY 764 ThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlyGluSerPheGly 783  
 DB 2464 ACACCAACAAATGATCTCAAGATCAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2523  
 QY 784 ThrSerGluLysPheGlnLysIleAsnGlnMetValLysAsnSerAspArgValIleLys 803  
 DB 2524 ACTTCTGCAAGTCTCCAGAAAAATAAATCAGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2583  
 QY 804 ArgSerAlaGlnGlySerAsnProProLysProLeuLysLysLeuArgPheAspIleGln 823  
 DB 2584 AGAATCTCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2643  
 QY 824 GlySerAspGluAlaAspGlySerLysHisIleProGlyGluSerLysPheGlnGlnLys 843  
 DB 2644 GATATGATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2703  
 QY 844 LeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetAsnAspSerMet 863

[illegible]







```

15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

```

```

10 2248 GACAAATATTGATGTTGTCATGATGGCATAGCAAAAGTGAATAATATAGACCTTAAA 2297
11 664 PheLysIleIleValThrAlaIleTyrAspLeuProHisAlaValGlnGluIleThrPheLys 683
12 2298 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2357
13 11CAAAATCAGTAAAGACATACAAAGATCTTCCATGCTGTCAGACAGACATTCAAA 2357
14 684 ArgValLeuIleLysGluGluIleTyrAspSerIleIleValPheTyrAspSerValPhe 703
15 2458 GTGCTTTTGAACAAATAGACAGATACATGATCTTATATAGTATCTTATTAACCTGGCTCTTC 2417
16 704 MetGluArgIleLysThrAsnIleLeuIleGlnIleTyrIleSerThrArgProThrLeuSer 723
17 2418 ATGAGACATGTGAATAATATTTTCTAGTATGCTTCCAGCAGGCGCTTATCTTGTGCA 2477
18 724 ProIleProHisIleProArgSerProIleTyrLysPheProSerSerProLeuArgIlePro 743
19 2478 CAAATATATCAATCTGCAAAATGCTTACCAAGTCTTCCATGCTTCCATGCTTCCATGCT 2547
20 744 GlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysIleSerGlyGlyLeuPro 763
21 2548 GACAGAAATATCTATATTTATATATGACAGATGCTTATATATATATATATATATATATAT 2597
22 764 ThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlyIleSerPheGly 783
23 2598 AATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2657
24 784 ThrSerGlnLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArgValIleLys 803
25 2658 AATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2717
26 804 ArgSerAlaIleLysSerAsnProProLysProLeuLysLysLeuArgPheAspIleGlu 823
27 2718 AGAAGATCTGACAGCAAGCAAGCTTCCATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2777
28 824 GlySerAspIleAlaAspGlySerLysHisLeuProGlyGluSerLysPheGlnGlnLys 843
29 2778 GATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2847
30 844 LeuAlaIleMetThrSerThrArgThrArgMetGlnLysGlnLysMetAsnAspSerMet 863
31 2848 GAGAGACAAATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2897
32 864 AspThrSerAsnLysGluLys 871
33 2898 CATATCTCAAAATCAAGCAACACAAA 2921

RESULTS
AR144797
LOCUS
DEFINITION Sequence 7 from patent US 6,210,939.
ACCESSION AR144797
VERSION AR144797.1 GI:15106564
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1. (bases 1 to 2995)
AUTHORS Gregory, R. L., Willis, K. N., and Manaval, D. G.
TITLE Recombinant adenoviral vector and methods of use
JOURNAL Patent: US 6,210,939 A 7 03 APR 2001;
FEATURES
Location/Qualifiers
1..2995
Source
BASE COUNT 975 a 617 c 594 g 810 t
ORIGIN
Alignment Statistics:
Procd. No. 1544 409 Length: 2995
Score: 4447.540 Matches: 878
Percent Similarity: 93.75% Conservative: 0
Best Local Similarity: 93.75% Mismatches: 1
Query Match: 98.70% Indels: 57
Gaps: 2

```

```

US-09-026-459a-49 (1-871) x AR144797 (1-2995)
11 MetProThrLysThrProArgLysThrAlaIleThrAlaIleAlaIleAlaIleAlaIleAlaIle 20
139 ATGGCGCGCAAAACCGCGCGGAAAAAAGCGCGGAAAGCGCGGAAAGCGCGGAAAGCGCGGAAAG 198
21 ProAlaProProProProProProProProProProProProProProProProProProProPro 40
199 CCGGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 258
41 AspLeuProLeuValArgLeuGluPheGluGluThrGluGluProAspPheThrAlaLeu 60
259 GAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 418
61 CysGlnIleLysIleLysIleProAspHisValArgGluArgAlaIleAlaIleAlaIleAlaIle 80
319 TGTCAGAAATTAAGATATACAGATCATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
81 ValSerSerValAspGlyValLeuGlyGlyTyrIleLeuIleLysIleLysIleLysIleLysIle 100
379 GTTTCATCTGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 438
101 IleCysIlePheIleAlaIleValAspLeuAspGluMetSerPheThrPheThrIleLeu 120
439 ATCTGTATCTTTATGTCAGCATGTGACTAGATGAGATGAGATGAGATGAGATGAGATGAGATG 498
121 GlnLysAsnIleLeuIleSerValHisLysPhePheAsnLeuLeuLysIleLeuAspThr 140
499 CAGAAAAACATATATATATATATATATATATATATATATATATATATATATATATATATATAT 558
141 SerThrLysValAspAsnAlaMetSerArgLeuLeuLysLysTyrAspValLeuPheAla 160
559 AGTACCAAGTCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 618
161 LeuPheSerLysLeuGluArgThrCysGluLeuIleTyrLeuThrGluProSerSerSer 180
619 CTCTTCAGCAATATGGAAGAGCATGTGAATATATATATATATATATATATATATATATATATATAT 678
181 IleSerThrGluIleAsnSerAlaLeuValLeuValSerIleIleThrLeuLeu 200
679 ATATCTACTGAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 748
201 AlaLysGlyValValLeuGlnMetGluAspAspLeuValIleSerPheGlnIleMetLeu 220
739 GCTAAAGCGGAAGTATATCAAAATGGAAGATGATGATGATGATGATGATGATGATGATGATGAT 798
221 CysValLeuAspTyrPheIleLysLeuSerProProMetLeuLeuLysIleLeuLeuLeuLeu 240
799 TGTGCTCTGCTATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 858
241 Thr-----GlySer----- 243
859 ACATCTTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 918
243 ----- 243
919 GCACGGATACGAAATCAATACGAAATCAATACGAAATCAATACGAAATCAATACGAAATCAATACG 978
243 ----- 243
979 CATGAAATGAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1038
244 AsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValIleLeuAsnLeuSerLysArg 263
1039 AATCTCTGAGCTTTTATGATCTTAATGATCTTAATGATCTTAATGATCTTAATGATCTTAATGATCT 1098
264 TyrGlnGluIleTyrLeuLysAsnLysAspSerAspAlaArgLeuIleLeuLeuLeuLeuLeuLeu 283
1099 TACGAGAAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1158
284 LysThrLeuGlnThrAspSerLysAspSerPheGluThrGluArgThrArgLysSer 303

```

```

1159 AAAACGTTTCAGACGCTATTTCTATAGACGCTTTTGAACACAGAGAAACACGAAAAAGT 1218
1204 ASNLEUASPGLUVALIASNVALILEPPTPTPHISLTHRPVALIARUHRVALMETASN 1278
1209 AACCTTCATGACAGAGGAGAAAGAAATTCCTCCACACATCCAGTTAGCAGCTGTATGAAC 1278
1214 ThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 1343
1219 ACTATCCAACTAATGATGATTTTAAATTCAGCAGTGATCAACCTTCAGAAAAATCTG 1338
1224 IleSerIyrPhaAsnAsnCysThrValAsnProLysGluSerIleLeuLysArgValLys 1363
1229 ATTTCATATTTAACTATGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
1234 AspIleCysIyrIlePheLysGluLysPheAlaValAlaValAlaValAlaValAlaVal 1383
1239 GATATGATATATATTTAAACAGAAAAATTTGCTAAAGCTGCGGACACAGGTTGCTCGAA 1458
1244 IleGlySerGlnAlaPyrLysLeuGlyValAlaLeuLysIyrIyrIyrIyrIyrIyrIyr 1403
1249 ATTGATATGACAGGATACAAACCTGAGTTCCTGTTATACCGAGTAATGGAATCCATG 1518
1254 LeuLysSerGluCysGluAlaGluSerIleGlnAsnPheSerLysLeuLeuAsnAspAsn 1423
1259 CTTAAATATGACAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1578
1264 IlePheHisMetSerLeuLeuAlaCysAlaLeuGluValAlaMetAlaIyrIyrSerArg 1443
1269 ATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1638
1274 SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProIleLeuAsnVal 1463
1279 ASATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1698
1284 LeuAsnLeuLysAlaPheAspPheLysValIleGluSerPheIleLysAlaIleLys 1483
1289 CTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1758
1294 AsnLeuThrArgGluMetIleLysHisLeuGluAlaCysGluHisArgIleMetGluSer 1503
1299 AACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1818
1304 LeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGlu 1523
1309 ATTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1878
1314 GlyProIleAspHisLeuGlnSerAlaCysProLeuAsnLeuProLeuGlnAsnAsnHis 1543
1319 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1938
1324 ThrAlaAlaAspMetIyrLeuSerProValArgSerProLysLysCysIyrSerThrThr 1563
1329 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1998
1334 ArgValAsnSerThrAlaAsnAlaSerThrAlaIleThrSerAlaPheLeuThrGlyLeu 1583
1339 CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2058
1344 ProLeuLysSerThrSerLeuSerLeuPheLysLysValIyrArgLeuAlaIyrLeu 1603
1349 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2118
1354 ArgLeuAsnThrLeuCysGluAlaGluLeuSerGluHisProGlnLeuGlnHisIleLeu 1624
1359 CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2178
1364 TrpThrLeuPheGlnIleThrLeuGlnAsnGluIyrGluLeuMetArgAspArgHisLeu 1643
1369 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2278
1374 AspGlnIleMetMetCysSerMetIyrGlyIleCysLysValLysAsnIleAspLysHis 1663
1379 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2298

```

```

QY 664 PheLysIleIleValThrAlaIyrLysAspLeuProIleAlaValGlnGlnThrPheLys 683
DB 2299 TCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2458
QY 684 ArgValLeuIleLysGlnGlnIyrAspSerIleIleValPheIyrAsnSerValPhe 703
DB 2359 CGTGTTCATCAAAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 2418
QY 704 MetGlnArgLeuLysThrAsnIleLeuGlnIyrAlaSerThrArgProProIleLeuSer 723
DB 2419 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2478
QY 724 PheIlePheHisIlePheCysSerProLysPhePheSerSerPheLeuArgAlaLeuPro 743
DB 2479 CCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2538
QY 744 GlyGlyAsnIleIyrIleSerProLysSerProLysIleSerGlyIleSerGlyIlePro 763
DB 2539 GAGGGAATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2598
QY 764 ThrPheThrIleGlnMetThrAlaSerArgIleLeuValSerIleGlyGluSerPheLys 783
DB 2599 AATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2658
QY 784 ThrSerGluLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArgValLeuLys 803
DB 2659 ACTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2718
QY 804 ArgSerAlaGlnGlySerAspProLysProLysLeuLysLeuAlaPheAspIleGln 823
DB 2719 AATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2778
QY 824 GlySerAspGluAlaAspLysSerLysHisLeuProGlyGluSerLysPheGlnGlnLys 843
DB 2779 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2838
QY 844 LeuAlaGlnMetThrSerThrArgThrArgMetGlnLysGlnLysMetAsnAspSerMet 863
DB 2839 CTGGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2898
QY 864 AspThrSerAspLysGluLys 871
DB 2899 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2922

RESULT 11
LOCUS HUMRBS 4740 bp mRNA linear FRI 12 JUL 1995
DEFINITION Human retinoblastoma susceptibility mRNA, complete cds.
ACCESSION M13400
VERSION M13400.1 GI:190958
KEYWORDS retinoblastoma susceptibility.
SOURCE Homo sapiens (clone: RA-11.5) foetus retina cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Mammalia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
Lee, W.H., Rockstein, R., Hong, F., Young, L.J., Shaw, J.Y., and Lee, P.Y.
Human retinoblastoma susceptibility gene: cloning, identification,
and sequence
Science 235 (4794), 1394-1399 (1987)
JOURNAL 87149066
MEDLINE 3923889
PUBMED 3923889
REFERENCE 2 (bases 1 to 480)
AUTHORS Lee, E.Y., Rockstein, R., Young, L.J., Lin, C.J., Rosenfeld, M.D., and
Lee, W.H.
TITLE Molecular mechanism of retinoblastoma gene inactivation in
retinoblastoma cell line Y79
Proc. Natl. Acad. Sci. U.S.A. 85 (16), 6017-6021 (1988)
JOURNAL 88320374
MEDLINE 3413073
PUBMED 3413073
COMMENT Draft entry and computer-readable copy of sequence in [1] kindly
provided by R. Rockstein, 27-APR-1987.

```



















Genature version 5.1.4  
Copyright (c) 1994 - 2003 CompuGen Ltd.

EM nucleic nucleic search, using SW model

Run on: January 16, 2003, 15:20:22 Search time 203.46 seconds  
(without alignments)  
18469.506 Million cell updates/sec

Filter: US-09-026-459a-38

Perfect score: 6224  
Sequence: 1 CGCGTATGCGCGCAAAAC.....AATGAGGATTATTCATAGT 3323

Scoring table: IDENTITY\_NP  
Gapop 10.0, Gapext 1.0

Searched: 16154066 steps, 809774476 residues

Total number of hits satisfying chosen parameters: 42408142

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listed first 45 summaries

Database: EST

1: em\_esth1\*

2: em\_esth2\*

3: em\_esth3\*

4: em\_esth4\*

5: em\_esth5\*

6: em\_esth6\*

7: em\_esth7\*

8: em\_esth8\*

9: qb\_est1\*

10: qb\_est2\*

11: qb\_est3\*

12: qb\_est4\*

13: qb\_est5\*

14: qb\_est6\*

15: em\_esth9\*

16: em\_esth10\*

17: qb\_qss1\*

18: em\_qss1\*

19: em\_qss2\*

20: em\_qss3\*

21: em\_qss4\*

22: em\_qss5\*

23: em\_qss6\*

24: em\_qss7\*

25: em\_qss8\*

26: em\_qss9\*

27: em\_qss10\*

28: em\_qss11\*

29: em\_qss12\*

30: em\_qss13\*

31: em\_qss14\*

32: em\_qss15\*

33: em\_qss16\*

34: em\_qss17\*

35: em\_qss18\*

36: em\_qss19\*

37: em\_qss20\*

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	910.2	27.4	1024	BM450031	BM450031 AGENCOURT
2	909.4	27.4	1051	BM451083	BM451083 AGENCOURT
3	856	25.8	1081	HQ439221	HQ439221 AGENCOURT
4	844.8	25.4	881	HQ220275	HQ220275 AGENCOURT
5	786	24.7	1001	BM475603	BM475603 AGENCOURT
6	752.8	22.7	1364	AK011246	AK011246 Mus muscu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	752.4	22.6	1122	13	BM450031	BM450031 AGENCOURT
8	726.4	21.8	727	9	AL592881	AL592881 DKFZp3140
9	720.6	21.7	889	14	HQ220275	HQ220275 AGENCOURT
10	699.6	21.1	823	12	HQ610661	HQ610661 G02611846
11	698.2	21.0	792	9	AL563757	AL563757 AL563757
12	692.2	20.8	767	12	HQ570456	HQ570456 G02591312
13	691.8	20.8	972	13	BM466194	BM466194 AGENCOURT
14	678	20.4	835	13	BM452724	BM452724 AGENCOURT
15	660	19.9	672	13	BM468359	BM468359 G03494340
16	645.6	19.4	797	10	HQ539278	HQ539278 G01064567
17	642.2	19.3	681	12	HQ616219	HQ616219 G02644717
18	617	18.6	618	13	BM464009	BM464009 G03410818
19	614	18.5	694	10	BE082846	BE082846 KC231064
20	605.2	18.2	623	10	AW581181	AW581181 JAL116917
21	605	18.2	605	9	AL599633	AL599633 DKFZp3140
22	596.4	17.9	922	12	HQ255343	HQ255343 G02364278
23	580	17.5	580	9	AL599195	AL599195 DKFZp3140
24	575.2	17.3	914	9	AA763411	AA763411 VW540410
25	566.4	17.0	588	10	AA768244	AA768244 CM510118
26	558	16.8	871	9	AA763405	AA763405 VW540410
27	530.2	16.0	598	10	AW582967	AW582967 JAL086087
28	530	15.9	539	9	AL692790	AL692790 DKFZp3141
29	528	15.9	572	9	AL120941	AL120941 DKFZp3140
30	516.8	15.6	597	12	HQ931246	HQ931246 M901015
31	513.2	15.4	594	10	AV715533	AV715533
32	512.2	15.4	734	12	BF144857	BF144857 G01790443
33	505.2	15.2	736	13	BM151568	BM151568 G02916108
34	493.8	14.9	525	9	AA592555	AA592555 Z1590041
35	486.4	14.6	584	12	HQ146050	HQ146050 G00860013
36	484.4	14.6	493	10	AW502887	AW502887 UT11F180
37	481.4	14.5	484	10	BE168095	BE168095 QV41051
38	480.2	14.5	659	13	BM211116	BM211116 G03035591
39	476.5	14.3	2083	13	BM456728	BM456728 AGENCOURT
40	474.4	14.3	734	13	BM526982	BM526982 G02927177
41	459.2	13.8	497	12	HQ934784	HQ934784 KC31094
42	457	13.8	457	9	AL093215	AL093215 G00440218
43	457	13.8	732	9	AL598766	AL598766 DKFZp3140
44	454.2	13.7	869	12	BE65749	BE65749 G01678131
45	447	13.5	447	9	AL082179	AL082179 G0360418

ALIGNMENTS

RESULT 1  
BM450031  
LOCUS BM450031 1023 bp mRNA linear EST 05 FEB 2002  
DEFINITION AGENCOURT\_6394544 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5128037  
5' mRNA sequence:  
BM450031  
BM450031.1 G118494071  
VERSION BM450031  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1023)  
AUTHORS NIH-MGC <http://mgi.nlm.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [quasr@mail.nih.gov](mailto:quasr@mail.nih.gov)  
Tissue procurement: AIC/DCU/bip  
cDNA library preparation: Life Technologies, Inc.  
cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LAM12204 row J column 22  
High quality sequence stop: 643  
Location/Qualifiers 1..1023

FEATURES

Source



















Pb	250	GAAGATATWKGATACATWTTTTAAAGCAAAATTCGTAAGATGTTGGGAAAGGGTGTGTGT	193
Qy	1092	CGAAATTCGATACACACCGCATACAAATCTGG-AGTTGCGTCTCTATATCCGAGTAAACGAA	1149
Pb	192	CGAAATTCGATACACGCAATCAACCTTGTTGCTTGTATTACCGAGTAATAGGA	183
Qy	1156	TGATATATTAAATCAGAGAGAAAGGATATTCATTCATTCAAAATTTTNGTAAGCTTCTAAT	1209
Pb	132	TGCAAGCTTATTCAGAAACAGACAGATATGCAATCAAAATTTTTCGACACATCTGCAAT	73

QY	1210	GACAA-CATTTCATATGCTTTTATTGGCGTGCGTCTCTTAGGTGTGAAT	1259
:			
B:	74	SA'AA-GVTTTTATATATAGTTTATTAGCGGTGGTGCTCCTAGGTTTGAAT	122
RESULT	12		
Locus	BG570456		
DEFINITION	hg570456 NM_025613123; NTH_MCF_77 Homo sapiens cDNA clone IMAGE471874 5'		
ACCESSION			
VERSION	hg570456		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		

REFERENCE:  
1. (BASES 1 TO 767)  
ATTHPS  
NIM M5C http://nim5c.aet.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1995)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cygabbz@mail.nih.gov  
Tissue Procurement: CLONTECH Laboratories, Inc.

```

cDNA Library Preparation: Clontech Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1568 row: h column: 03
High quality sequence step: 726.
Location/Qualifiers
1. 767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:4717874"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (TI phase-resistant)"
/note="organ: lung; Vector: pINR-1.1B (Clontech); Site:
Sfil (ggccgcctccgcc). Site2: Sfil (agccattatgcc); 5'
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAAGGCACTATATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGAGCGAGCATG-dT(30)BN-3' (where B =
C, G or A, C, G, or T). Average insert size
kb (range: 0.5-4.9 kb). 12,715 colonies contained insert
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto
CA). Note: this is a NIH_MGC Library."
BASE 690NT 233 a 125 c 120 g 279 t
ORIGIN
Query Match 20.4%; Score 692.2; DH 12; Length 767;
Percent local similarity 97.1%; Pred. No. 20-125;
Matches 747; Conservative 0; Mismatches 18; Indels 4; Gaps
QY 2531 GCATGATATCTCAACCAAGGAGAGCAATAGAGATCTCAAGACCTTGGAGACATGCG 2560
|||||
Db 1 GCATGATATCTCAACCAAGGAGAGCAATAGAGATCTCAAGACCTTGGAGACATGCG 60
QY 2591 TACACCTCTGGGATTCATTGCTCTCTCAAGATGTCAGTACTGATATAACTTTGCAGAGTTCTGTT 2650
|||||

```

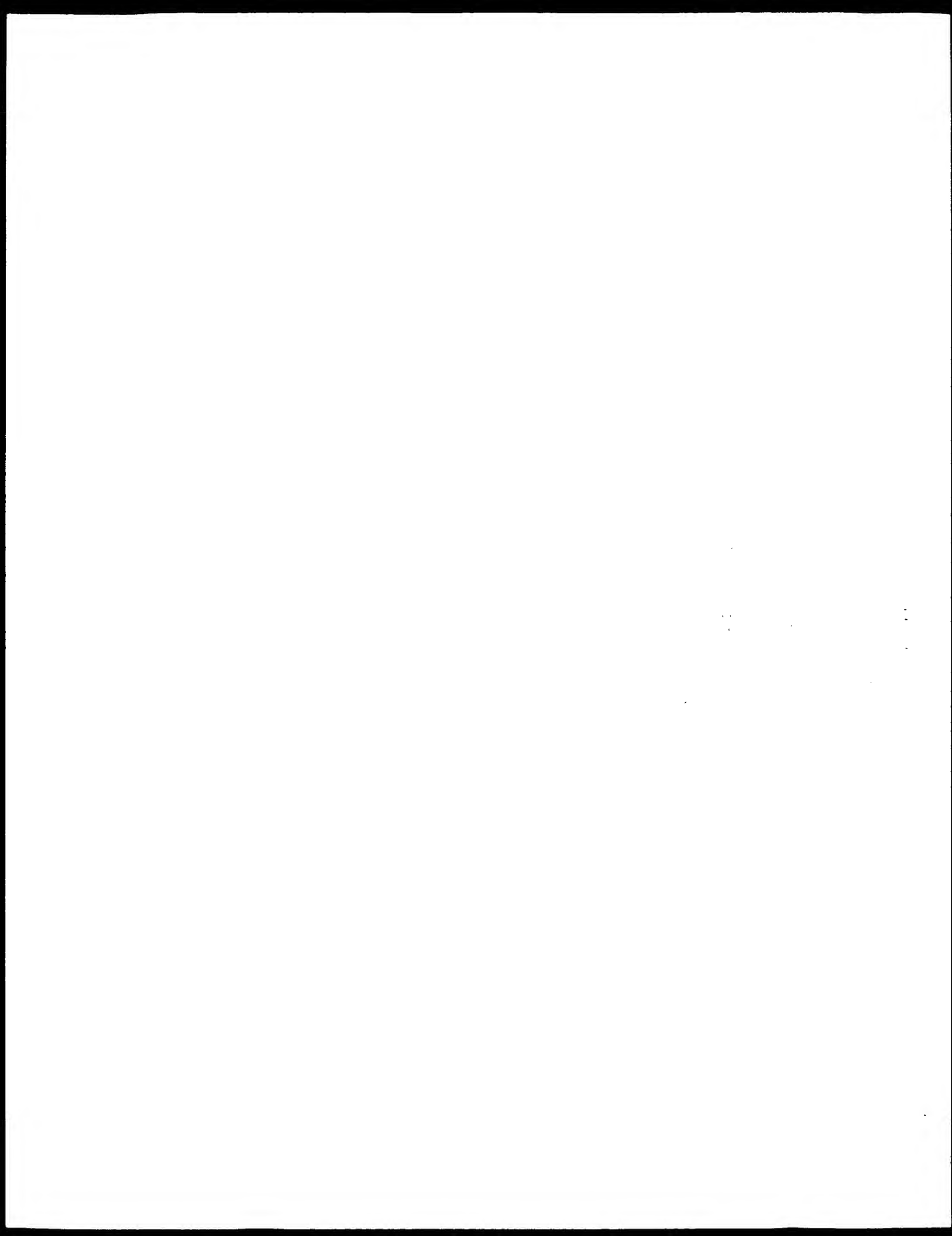






2249	TTACTGCAAGATGAAAGAAATTATATATCAATTTGGTGAATCATTTCTGGGATCTCTGGAAAT	2308
10b		
61	TGATCTCAAGATCAAGAAATCTTAAATATCAATTTGCTGAATCATTCGGACATCTTCAGAACT	120
10b		
2309	TTCAAAAAATAATCATATGATATTAACACAGAAATGTTGTGCTCAAAAATAATGTTGAAG	2368
10b		
121	TTCAAAAAATAATCATATGATATTAACACAGAAATGTTGTGCTCAAAAATAATGTTGAAG	180
10b		
2369	GAAATTAATCTCTCTAAATATCTGAAAAATTAAGCTTTTGATATTAAGGATACAGATGAAG	2428
10b		
181	GAAATTAATCTCTCTAAATCAATGAAAAAATAATAGTTTGTATATTGAAGGATCATAGATGAAG	240
10b		
2429	CAGATGAAGTAAATATATCTCCAGAGAGTCCAAATTTCCAGCAGAACTGGTAGAATCA	2488
10b		
241	CAGATGAAGTAAATATATCTCTCCAGAGATCCAAATTTCCAGCAGAACTGGTAGAATCA	300
10b		
2489	CTTCTACTAGAACATGAATGCAAAAAGAGAAAATTAATATAGCATGATATGCTCAAGA	2548
10b		
301	CTTCTACTAGAACATGAATGCAAAAAGAGAAAATTAATATAGCATGATATGCTCAAGA	360
10b		
2549	AGACACAAAAAGATGATCTTACATACCTTGGACACATCTGACACCTCTGGATTCATTT	2608
10b		
361	AGACACAAAAAGATGATCTTACATACCTTGGACACATCTGACACCTCTGGATTCATTT	420
10b		
2609	GTCTCTCAATGATGATGATATAATCTTCCACATCTCTGTTATGCGACATTTTAAT	2668
10b		
421	GTCTCTCAATGATGATGATATAATCTTCCACATCTCTGTTATGCGACATTTTAAT	480
10b		
2669	CTTCAATCTCTTTTGTGATATAAAATGTGCAATATGCAATTTGTTTGGTGAATCTTAAGC	2728
10b		
481	CTTCAATCTCTTTTGTGATATAAAATGTGCAATATGCAATTTGTTTGGTGAATCTTAAGC	540
10b		
2729	CAATTCAGAAAGTATGATATATTTATACAGATTCAGAAATATCTCTGTTGAATCTCTGC	2787
10b		
541	CAATTCAGAAAGTATGATATATTTATATACAGATTCAGAAATATCTCTGTTGAATCTCTGC	600
10b		
2788	CATTTAAAAAATTTGTATGCAATATTTTCTCTCTTCCAAAGTAAAAATTTCTGTGTCTTTATGG	2847
10b		
601	CATTTAAAAAATTTGTATGCAATATTTTCTCTCTTCCAAAGTAAAAATTTCTGTGTCTTTATGG	660
10b		
2848	ATAGTAAATAATG	2859
10b		
661	ATAGTAAATAATG	672

Search completed: January 18, 2003, 04:08:48  
Job time : 2939.95 secs













```

FEATURES
Source
Location/Qualifiers
1..835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_72"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-Sport6; Site: 1. NotI;
Site: 2. SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 254 a 210 c 137 q 233 t 1 others
RESULT 6
HM475603
LOCUS
DEFINITION
5', mRNA sequence.
ACCESSION
BM475603
VERSION
BM475603.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota, Metazoa, Chordata, Craniata; Vertebrata; Euteleostomi;
Mammalia, Eutheria, Primates, Catartini; Homidae; Homo.
REFERENCE
1 (bases 1 to 1001)
AUTHORS
NIH-MGC http://mgs.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rstraus@nsl.nih.gov
Tissue procurement: ATCC
cDNA library preparation: Life Technologies, Inc.
cDNA library arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: HAM12326 Row: f Column: 04
High quality sequence start: 17
High quality sequence stop: 734.
FEATURES
source
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-Sport6; Site: 1. NotI;
Site: 2. SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC library."
BASE COUNT 349 a 178 c 186 g 288 t
ORIGIN
Alignment Scores:
Pred. No.: 1 99e-104 Length: 1001
Score: 1152.00 Matches: 242
Percent Similarity: 90.43% Conservative: 13
Best local Similarity: 85.82% Mismatches: 15
Query Match: 25.87% Indels: 12
DB: 13 Gaps: 4
US-09-026-459a-43 (1-859) x BM475603 (1-1001)
QY 93 GlnLysLysLysGlnSerTrpClyLeuCysLeuPheLeuAlaValAspLeuValCys 112
DB 132 GAAAGACACCGTGAACATTT-----ATAATTTTGACACCAACCCAGCTTCGATA 179
QY 113 SerThrGluLeuAsnSerAlaLeuValLeuLysValSerTrpLeuPheLeuAla 132
DB 180 TCTACTGAAATTAATTCGCAATTCGCTAAAGGTTCTTGGATCACAATTTATACCT 239
QY 133 LysClyGluValLeuGlnMetGluAspLeuValLeuSerPheGlnLeuMetLeuCys 152
DB 240 AAGGGAAGTATTACAAATGGAAGATGATCTGTATTCATTTCATTTCATTTCATTTC 299

```







with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ193 vector. Library is normalized. Library was constructed by Hento Soares and M. Fatima Ronaldo.

BASE COUNT 257 a 178 c 217 g 261 t 1 others

Alignment Scores: 6,160-95 Length: 914  
Pred. No.: 1056-50 Matches: 234  
Percent Similarity: 82.14% Conservativity: 19  
Best Local Similarity: 75.97% Mismatches: 53  
Query Match: 23.73% Indels: 6  
DB: 9 Gaps: 2

US-09-026-459a-43 (1-859) x AA764411 (1-914)

```

QY 267 PheLeuAspHisAspLysThrLeuGlnThrAspSerLeuAspSerPheGluThrGlnArq 286
DB 914 TTTTGCAATGACAGAAATAAGCCCTCCAGCTCT--TCCTTAGCCCTTTTAAAGCGCGAG 857
QY 287 ThrProAlaLysSerAsnLeuAspGluGluValAsnValIleProProHisThrProVal 306
DB 856 ACCGACGGAAACACACCTTGT-GAAGCGGTAAACAGAGTTAAGCGGTAAAGCGGT 798
QY 407 ArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGln 326
DB 797 TGGGCTGTTTGGAAACGACGACCAATTTATGTCATTTAAACCTCCGCCCAATCACCGC 738
QY 427 ProSerLeuAsnLeuLeuLeuSerTyrPheAsnAsnCysThrValAsnProLysGluSerIle 346
DB 737 GCGTCAGAAATTCGTGT-TGGCTCTGTAATATGGCGCGGGAAGAAAGCAAGAAATTCGC 679
QY 347 LeuLysArgValLysAspIleGlyTyrIlePheGlySerGluSerPheAlaValGly 366
DB 678 CTAAAGCGGT-ANGAGGTGGACATGCTTTAAGAGAAATTTCTA---ACGCTGTGG 623
QY 367 GluGlyCysValGluIleGlySerGlnArgTyrLysLeuGlyValArgLeuTyrTyrArg 386
DB 622 CCAAGCTGTGTGTATATACATGATACAGCAATATATATATATATATATATATATATAT 563
QY 387 ValMetGluSerMetLeuLysSerGluGluGluLeuSerIleGluAsnPheSerLys 406
DB 562 GTCATGGAAATCCATGCTATACATGACAGAGAACGTTTCCATTCACAAATTTAGCAAA 503
QY 407 LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuValValMet 426
DB 502 CTCTTAATGCAACATCTTCATATGTTTACTGCTGTGTGTGTGTGTGTGTGTGTGTGT 443
QY 427 AlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro 446
DB 442 GTATGCTATAGCAAGATATATGACCATCTTGTGATTCGGACACATTCCTCTCCCG 383
QY 447 TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPhe 466
DB 382 TGAATTCGACACCTACCTAATATATATATATATATATATATATATATATATATATAT 323
QY 467 IleLysAlaGluThrLysLeuThrArgGluMetIleLysHisLeuGluArgCysGluHis 486
DB 422 ATCAAAATGCAAGCAACTGCAAGAGAAATGATATAAAATTTTACAAAGATGTGAGCAT 263
QY 487 ArgIleMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuLysGln 506
DB 262 CCAATCATGCAATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 203
QY 507 SerLysAspArgGluThrProThrAspHisLeuGluSerAlaCysProLeuAsnLeuPro 526
DB 202 TCAAGCAATGCAAGCAACT---GATAAGCTGTGAACCTGCTGTGCTGTGCTGTGCTGTG 146
QY 527 LeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys 546
DB 146 CTGACAGGTAAACCATGTCAGAGAGATATATATATATATATATATATATATATATATAT 86

```

```

QY 547 LysGlySerThrThrArqValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAla 566
DB 85 AGAACTTCCACTACACGCTGTAATTCCTGCGCAAAATACACACACACACACACACAC 26
QY 567 PheGluThrGlnLysProLeuLys 574
DB 25 TTCATACACTCGAAGGCAATTGAAA 2
RESULT 10
AL563757/c 792 bp mRNA linear EST 16 FEB 2001
LOCUS AL563757 LTI_NFL301_NBC4 Homo sapiens cDNA clone (S014007YA05) 4
DEFINITION prime, mRNA sequence.
ACCESSION AL563757
VERSION AL563757.1 GI:12913464
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 792)
AUTHORS Li, W.H., Gruber, C., Jossee, J., and Pollayes, D.
TITLE Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
HP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
source
1..792
organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS000007YA05"
/clone_lib="LTI_NFL301_NBC4"
/sex="male"
/lissup_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT 246 a 148 c 117 g 255 t 26 others
ORIGIN

```

Alignment Scores:  
Pred. No.: 1,120-93 Length: 792  
Score: 1043.00 Matches: 217  
Percent Similarity: 84.53% Conservativity: 7  
Best Local Similarity: 81.89% Mismatches: 38  
Query Match: 23.42% Indels: 3  
DB: 9 Gaps: 0

US-09-026-459a-43 (1-859) x AL563757 (1-792)

```

QY 171 LysThrAlaValIleProLleAsnGlySerProArgThrProArgArqGlyGlnAsnArq 190
DB 791 AAAACACGCTTCTATACCCATTAATGCTTCCACCTCCAGCAATCCAGCGGACGTCAGAAC 732
QY 191 SerAlaArgIleAlaLysGlnLeuGluAsnAspThrArgIleIleGluValIleCysLys 210
DB 731 AGTCCACGCGTATGACACACAAATAGAAATATGATACAGAAATATTCAGATTCATCTAAA 672
QY 211 GluHisGluCysAsnIleAspGluValLysAsnValTyrIleLysAsnIleIleProLe 230
DB 671 GAAATATGATGTATATATAATGAGCTGCAAAATGTTTWTTCAAAAATTTTATATATTTT 612

```

```

QY 241 MetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProValValGluValAsnLeuSerLys 250
DB 611 ATCAATCTCTCTGGCACTCTTACACATCTATGCACTCTCAGAGCTTCAAAATTTCTTAA 552
QY 251 ATGATCTCTCTCTGGCACTCTTACACATCTATGCACTCTCAGAGCTTCAAAATTTCTTAA 270
DB 611 ATCAATCTCTCTGGCACTCTTACACATCTATGCACTCTCAGAGCTTCAAAATTTCTTAA 552
QY 271 AspLysThrLeuSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 290
DB 491 GATAAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 432
QY 291 SerAsnLeuAspGluGluValAsnValIleProProHisThrProValArgThrValMet 310
DB 431 AGTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 372
QY 311 AsnThrIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 330
DB 471 GATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 312
QY 331 LeuLeuSerThrLeuAsnAspCysThrValAsnProLysGluSerIleLeuLysArgVal 350
DB 411 YGTHHCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 252
QY 351 LysAspIleGlyThrIlePheLysGluLysPheAlaLysAlaValGlyGlyGlyCysVal 370
DB 251 AAGCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 192
QY 371 GluIleGlySerGlnArgThrLysLeuGly--ValArgLeuThrThrArgValMetGlu 390
DB 191 GAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 132
QY 391 GlnMetLeuLysSerGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 409
DB 131 CCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 72
QY 410 AspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrTyr 429
DB 71 ACAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 12
QY 430 SerArgSer 432
DB 11 GCATTAAT 3

RESULTS 11
AL599633
DEFINITION DKF2033P092.1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION AL599633
VERSION AL599633.1 GI:15162921
SOURCE EST
ORGANISM human
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Primates; Catarrhini; Hominoidea; Hominidae; Homo;
Homo sapiens (bases 1 to 605)
AUTHORS Koehler, K., Beyer, A., Mewes, H. W., Weil, B., and Wiemann, S.
TITLE EST (Koehler, K., Beyer, A., Mewes, H. W., Weil, B., and Wiemann, S.)
JOURNAL Unpublished (1999)
COMMENT Contact: Koehler K
MIPS Am Kleinferspitze 18a D-82152 Martinsried, Germany
this is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No 5' sequence available.
This clone (DKF2033P092) is available at the EMBL in Berlin.
Please contact the RZPD, Ressourcenzentrum, Heuberweg 6, 14059
Berlin-Charlottenburg, GERMANY. Email: clone-arp@rzpd.de

```

```

FEATURES
    source          location/qualifiers
    1..605
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="313 (synonym: hlcc2)"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Vector, pTribex2; Site_1: SfilA; Site_2: SfilR;
    cDNA collection"
BASE COUNT      228 a   32 c   105 g   179 t
ORIGIN
Alignment Scores:
    4 016-93      length:      605
    Score:        1037.00      Matches:      201
    Percent Similarity: 100.00%      Conservative: 0
    Best Local Similarity: 100.00%      Mismatches: 0
    Query Match:    23.29%      Indels:      0
    DB:              Gaps:      0
US-09-026-459a-43 (1-859) x AL599633 (1-605)
QY 206 GluValLeuCysLysLysIleGluCysAsnIleAspGluValLysAsnValTyrPheLys 225
DB 1 GAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
QY 226 AspPheIleProPheMetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluVal 245
DB 61 AATTTTATACTTTTATTAATTTCTTGAATTTCTTGAATTTCTTGAATTTCTTGAATTT 120
QY 246 GluAsnLeuSerLysArgTyrGluGluIleTyrLeuLysAsnLysAspLeuAspAlaArg 265
DB 121 GAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 266 LeuPheLeuAspHisAspLysThrLeuGlnThrAspSerIleAspSerPheGluThrGln 285
DB 181 TTATTTTGGATCATGATAAACTCTTCTGAGCTATTTCTATAGACACTTTTGAACATAG 240
QY 286 ArgThrProArgLysSerAsnLeuAspGluGluValAsnValIleProProHisThrPro 305
DB 241 AGAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 306 ValArgThrValMetAsnThrIleGlnGlnLeuMetIleLeuAsnSerAlaSerAsp 325
DB 301 GTTAGGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 326 GlnProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProLysGluSer 345
DB 361 CAATTTTCAAAATTTCTATTTTCTATTTTCTATTTTCTATTTTCTATTTTCTATTTT 420
QY 346 IleLeuLysArgValLysAspIleGlyTyrIlePheGlySerGlyLysPheAlaLysVal 365
DB 421 ATACTCAAAAGAGTCAAGCATATAGATACATCTTTAAAGAGAGAGAGAGAGAGAGAG 480
QY 366 GlyGlnGlyCysValGluIleGlySerGlnArgTyrLysLeuGlyValArgLeuTyrTyr 385
DB 481 AGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 386 ArgValMetGluSerMetLeuLysSerGluGluGluArgLeuSerIleGluAsnPheSer 405
DB 541 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
QY 406 Lys 406
DB 601 AAA 603
RESULTS 12
LOCUS BE082846
DEFINITION BE082846
ACCESSION BE082846
VERSION BE082846.1 GI:8473151

```

```

RESULTS 12
LOCUS BE082846
DEFINITION BE082846
ACCESSION BE082846
VERSION BE082846.1 GI:8473151

```

# KEYWORDS SOURCE ORGANISM

EST.  
human.  
Homo sapiens

# REFERENCE AUTHORS

1 (bases 1 to 694)  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bello, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Donquele, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J.

# TITLE

Shedding sequencing of the human transcriptome with ORF expressed sequence tags

# JOURNAL MEDLINE COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
2002664  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55 11 2704922  
Fax: +55 11 2707001  
Email: asimpson@ludwig.org.br

# FEATURES SOURCE

1..694  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="H10642"  
/note="organ: breast; Vector: puelB; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORFESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the pOR 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

# BASE COUNT GCTGTC

211 A 129 C 127 G 226 T 1 others

# Alignment Scores: Prod. No.: Score: Percent Similarity: Best Local Similarity: Query Match: DB:

1..844..92 Length:  
1010..00 Matches:  
96.9% Conservative:  
94.95% Mismatches:  
24.14% Indels:  
10 Gaps:

# US 09 026 459A 43 (1-694) X BE082846 (1-694)

QY 437 AspCysThrValAsnProLysGluSerLeuLysAcqValLysAspIleGlyTyrIle 356  
Db 676 CACACACACATGACAAAGAAAGTATACAGACAGACAGTCAAGCATATACATACATC 618  
QY 657 PheLysGluLysPheAlaValAlaValGlyGlnGlyCysValGlnIleGlySerGlnArg 376  
Db 617 TTTAAACAGAAATTCGTAGAGTGTGGGACAGNNTGTGTGCAAAATTCATCAGACAGCA 558  
QY 477 TTTTysGluGlyValAlaGlnIleGlyTyrIleValMetGluSerMetLeuLysSerGluGlu 396  
Db 557 TAAAAATTCACATTCGTGTGTAATACACAGTAAATTCATCAGTAAATTCATCAGTAA 498  
QY 497 GluArgLeuSerIleGlnAsnProLysLysLeuLysAsnAspAsnIlePheIleMetSer 416  
Db 497 GAGGATTAATCCATTCATCAAAATTTAGCAACATTCGATGACAAACATTTTCATATGCT 438

# QY 417

LeuLeuAlaCysAlaLeuGluValValMetAlaIleThrIleSerArgSerThrSerGlnAsn 436  
Db 437 TTAATGGGCGGCGCTTGGAGTTGTAATGGGCAATATATAGCAGAGATGATATCTAGAAAT 378

# QY 437

LeuAspSerGlyThrAspLeuSerPheProTrrIleLeuAsnValLeuAsnLeuLysAla 456  
Db 377 CTTCATCTCGGACACATATAGCTTTCCCATGCAATCTCAATGCTTAAATTTAAAGCC 318

# QY 457

PheAspPheTyrLysValIleGluSerPheIleLeuAlaGlnGlyAspIleThrAlaGlu 476  
Db 317 TTGGATTTTACAAAGTATGCAAGCTTTATCAAAAGAGAGAGCACTTCATAGACAA 258

# QY 477

MetIleLysHisLeuGluArgCysGluHisArgIleMetGluSerLeuAlaIlePheSer 496  
Db 257 ATGATAAAATATTAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATG 198

# QY 497

AspSerProLeuPheAspLeuIleLysGluSerPheAspIleGlnIleGlnIleGlnIle 516  
Db 197 GATTCACCTTTATTCATGCTTTATTAACAAATCAAGAGATCAGACAGCAATCAATC 138

# QY 517

LeuGluSerAlaCysProLeuAsnLeu-ProLeuGlnAsnAsnIleThrAlaAlaAspMe 536  
Db 137 CTGGAATCTGCTGTCTTAAATGCTGTTCTGCAAAATTAAGAGCTTATCAATGCTAA 78

# QY 536

TyrLeuSerProValArgSerProLysLysLysSerIleThrAlaVal 554  
Db 77 GTATCTTCTGCTGTAATGCTGCAAAATTAAGAGCTTATCAATGCTAA 26

# RESULT 13 LOCUS DEFINITION

AW583181 623 bp mRNA linear EST 13-MAR-2002  
1al1b09.y1 Human Pancreatic Islets Homo sapiens cDNA clone  
IMAGE:5638049 5' similar to qh:M15400 RETIN-BLASTOMA-ASSOCIATED  
PROTEIN (HUMAN); qh:M26391 Mouse retinoblastoma susceptibility  
protein (MCUSE); mRNA sequence.

# ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AW583181  
AW583181.1 GI:7260125  
EST.  
human.  
Homo sapiens

# REFERENCE AUTHORS

1 (bases 1 to 623)  
Wyllie, T., Meadows, A., Clifton, S., Hillier, S., Maria, M., Pape, D.,  
Wyllie, T., Martin, J., Bilstain, A., Schmitt, A., Theising, B., Rittner,  
R., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., Merdun, R., Cole,  
R., Tsakareishvili, R., Williams, T., Jackson, Y., and Lowers, Y.  
WashU-Harvard Pancreas EST Project  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu

# TITLE JOURNAL COMMENT

Libraries were constructed by Dr. Douglas Melton  
DNA sequencing by: Washington University Genome Sequencing Center  
For information on obtaining a clone please contact: Juliana Brown  
(brownefas.harvard.edu)  
This sequence now available from the IMAGE consortium, for clone  
orders contact: info@image.llnl.gov  
Seq primer: -40RP from gibco  
High quality sequence stop: 440.  
Location/Qualifiers  
1..623  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:5638049"  
/tissue\_type="Islets of Langerhans"  
/dev\_stage="Adult"

# FEATURES SOURCE

1..623  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:5638049"  
/tissue\_type="Islets of Langerhans"  
/dev\_stage="Adult"

```

/lab_host="DH10B"
/Note="Organ: pancreas, Vector: pSP6K11, Site: Not i,
Site 2: Sal I; Library constructed using SuperScript
Plasmid Library kit (Life Technologies). cDNA made by
oligo-dT priming. Size-selected by column fractionation;
average insert size 1.08 kb. Primary library,
unamplified."

```

```

BASE COUNT      234 a  93 c  109 g  186 t      1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3,950-91      Length:      623
Score:          1016.00      Matches:      201
Percent Similarity: 98.07%      Conservative:  2
Best Local Similarity: 97.10%      Mismatches:   3
Query Match:      22.82%      Indels:       1
DB:              10          Gaps:         0

```

```

US-09-026-459a-43 (1-859) x AW583181 (1-623)

```

```

QY 202 ThrAaTleTleTValLeuGysTysGThSsGlyLysAsnValValLysAsn 221
|||||
Db 3 ACAGAAATATTGCAAGTCTGTAAACAAACATCAATATATATAGATCAAGTGAATAAT 62
|||||
QY 222 ValTyrPheLysAsnPhelIleProPheMetAsn-SerLeuGlyLeuValThrSerAsnGI 241
|||||
Db 63 GTTATTTCAGAAATTTTATAGCTTTTATGGAATTCCTTGGACTTGTAAATCATATGG 122
|||||
QY 241 yLeuProGluValGluAsnLeuSerLysArgTyrGluGluIleTyrLeuLysAsnLysAs 261
|||||
Db 123 ACATTCACAGGTGTGAAGATCTTTCGAAAGATATCAAGCAAAATTTATCTTAAATAAGA 182
|||||
QY 261 pLeuAspAlaArgLeuPheLeuAspHisAspLysThrLeuGlnThrAspSerIleAspSe 281
|||||
Db 183 TCTAGATCAAGATATTATTTCGATCATGATAAATCTTTAGACTGATTCATAGACAG 242
|||||
QY 281 rPheGlnThrGlnArgThrProArgLysSerAsnLeuAspGluGluValAsnValIlePr 301
|||||
Db 243 TTTTGAAGAACACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 302
|||||
QY 401 oProHisThrProValArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAs 321
|||||
Db 403 TCCACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 362
|||||
QY 421 nSerAlaSerAspGlnProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAs 341
|||||
Db 463 TTCTAGCAATGTATCAACCTTTCAAAAATCTGATTTCTATTTTAAACAACTCAACAGTAA 422
|||||
QY 341 nProGlnSerThrProValArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeu 361
|||||
Db 423 TTTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 482
|||||
QY 361 eAlaLysAlaValGlyGlnGlyCysValGluIleGlySerGlnArgTyrLysLeuGlyVa 381
|||||
Db 483 TCTTAAAGCTCTCGACACAGCTTCTGTCGCAAAATTCATCAGCGCATACAACTGGAGT 542
|||||
QY 481 lArAsnTyrTyrArgValMetGlnLysMetLeuLysSerGluGluValArgLeuSerTl 401
|||||
Db 543 TCTTGTGTATTACAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 602
|||||
QY 401 eGlnAsnPhoSerLysLeu 407
|||||
Db 603 TCAGAAATTTTATGCAAACTT 621

```

```

RESULT 14
AA7634485.1
LOCUS
DEFINITION
vW54h01 r1 Soares mammary gland NMIMC Mus musculus cDNA clone
IMAGE:1247665.5' Similar to gb-M5466 FETINBLASTOMA ASSOCIATED
PROTEIN (HUMAN); gb-M26341 Mouse r1FetInblastoma susceptibility
protein (MOUSE); mRNA sequence.
AA7634485
AA7634485.1 GI:2813232

```

# KEYWORDS SOURCE ORGANISM

## REFERENCE AUTHORS

## TITLE JOURNAL COMMENT

```

EST.
house mouse.
Mus musculus.
Eukaryota: Metazoa: Chordata: Vertebrata, Euteleostomi:
Mammalia, Eutheria: Rodentia, Sciurognathi, Muridae: Murinae: Mus.
1 (bases 1 to 871)
Matta,M., Hillier,L., Allen,M., Howies,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kuvshinov,T., Lacy,M., Lee,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Matta M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty free through LNC. Contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI-661353
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 495.
Location/Qualifiers
1..871
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1247665"
/clone_lib="Soares_mammary_gland_NMIMC"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Benito Soares and M. Fatima Ronaldo."

```

```

BASE COUNT      242 a  172 c  204 g  251 t
ORIGIN

```

## FEATURES source

## Alignment Scores:

```

Pred. No.:      1,490-90      Length:      871
Score:          1013.00      Matches:      216
Percent Similarity: 90.34%      Conservative:  17
Best Local Similarity: 74.48%      Mismatches:   53
Query Match:      22.75%      Indels:       5
DB:              9          Gaps:         2

```

```

US-09-026-459a-43 (1-859) x AA763485 (1-871)

```

```

QY 287 ThrProArgLysSerAsn-LeuAspGlnGluValAsnValIleProHisThrProV 306
|||||
Db 860 AGGCAAGGAAAAAAACAGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804
|||||
QY 306 aLArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspG 326
|||||
Db 803 TTGG-CCCGTTTTTCATCTACTCCACCATTAGGGGGGTTTTTAATTTCCCAAGTATCA 745
|||||
QY 326 InProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProLysGluSerI 346
|||||
Db 744 GAGCAACGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
|||||
QY 346 lLeuLysArgValLysAspIleGlyTyrIlePheLysGluLysPheAlaLysAlaValC 366
|||||
Db 684 TCTAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 625
|||||
QY 366 lYcGlnGlyCysValGluIleGlySerGlnArgTyrLysGlnValArgTyrLysTyrA 386
|||||

```

```

624 GTTACGCTTGGTGGACATACAGGATATACACATTCAGGACGCGCATTTATTACC 565
QY rValMetGluSerMetLeuTyrSerGlnGluArqLeuSerIleThrAspPheSerL 406
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
664 GTGGATCATATGATGATATATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 505
QY ystLeuLeuAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValuIM 426
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
604 CATCTTAATATGACACATCTTTCATATGCTCTATCTGCGCTGCTCTGAAGTTGTA 445
QY eAlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlnTyrAspLeuSerPheP 446
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 GATATATATATATATATATATATATATATATATATATATATATATATATATAT 385
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 466
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
684 GATGATCTGAGAGATATATATATATATATATATATATATATATATATATATAT 425
QY hTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 486
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
624 TTATATATATATATATATATATATATATATATATATATATATATATATATAT 245
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 506
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
664 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 205
QY hTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 526
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
604 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 148
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 546
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
624 rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 566
QY ystLeuLeuAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValuIM 506
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 88
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 88
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
624 ystLeuLeuAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValuIM 566
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 566
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 28
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 28
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
624 rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 574
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 574
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 2
QY rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerP 2
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

BASE COUNT 307 a 214 c 204 g 197 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1110-89 Length: 922  
Score: 1004.00 Matches: 224  
Percent Similarity: 90.51% Conservatives: 5  
Best Local Similarity: 88.54% Mismatches: 21  
Query Match: 22.55% Indels: 8  
DB: 12 Gaps: 0  
SS-99-026-459A-43 (1 853) x B625454 (1 922)  
QY 610 IleIleTrpThrLeuPheGlnHisThrLeuGlnAsnGluTyrGlnLeuLeuMetValAsnAr 629  
DB 1 ATCATCTGGACCCCTTTCAGACACACACCCCTGACAAACAGCAAGCAATTCATCAGATACAG 60  
QY 629 rHisLeuAspIleLeuMetMetCysSerMetTyrGlnTyrValTyrValTyrValTyrVal 649  
DB 61 GCATTG-GACCAAAATTATGATGTGTTCATGTATGATGATGATGATGATGATGATGATGATGAT 119  
QY 649 pLeuLysPheLysIleLeuValTyrAlaTyrLysAspLeuThrAlaValGlnGlnGlnGln 669  
DB 120 GCTTAAATTCAAATATCAT-GTAAACAGCATATACAGATATCTCTCTCTCTCTCTCTCTCTCT 178  
QY 669 rPheLysArgValLeuIleLysIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 689  
DB 179 ATTCAACGGCTCTGCAACAAACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 247  
QY 689 rValPheMetGlnArqLeuLysThrAsnIleLeuGlnIleTyrAlaSerThrArqProProth 709  
DB 248 GGTCTTCATGACAGACATGACAAACAAATAT TTTGATATGTTTTCATGACAGACAGACAG 296  
QY 709 rLeuSerProIleProHisIleProArqSerProTyrLysPheProSerSerProLeuAr 729  
DB 297 CTTCACCAATACCTCACAATCTCTGAAAGCTTACAGTTTCCTAGTTTCAGTTTCAGTTTCAG 456  
QY 729 rIleProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 749  
DB 357 GATTCCTGGAGGAGCAAT 416  
QY 749 rLeuProThrProThrLysMetThrProArqSerArqIleLeuValSerIleGlnGlnGln 769  
DB 417 TCTGTCAACACCAACAAACAAATATGATCTGAAAGCTTACAGTTTCCTAGTTTCAGTTTCAG 476  
QY 769 rPheGlnTyrSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 789  
DB 477 ATTGAGGATCTCTGAGAGCTTCCAGAAATATATATATATATATATATATATATATATATAT 546  
QY 789 rLeuLysArgSerAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 809  
DB 537 GCTCAAAAGAGAGTCTGCAAGGAG 595  
QY 809 rIleGln 828  
DB 596 TATTGAAAGAT 655  
QY 829 Gln 848  
DB 656 CAGGCAAAATCTGGGCAAAATGATCTTATTCGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 715  
QY 849 AspSerMetAspThrSerAsnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 859  
DB 716 GGAAGCTGGCTATCTCTCAGACAG 748

```

/cclone="IMAGE:4472611"
/cclone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="organ: liver; Vector: pCMV Sport6; Site: 1; Not 1;
Site 2; Salt; cloned unidirectionally; original primed;
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by life technologies.
Note: this is a NIH MGC library."

```

REFERENCE:
AUTHORS:
TITLE:
JOURNAL:
COMMENT:

Contact: Robert Strausberg, Ph.D.  
Email: rstraus@nhi.nih.gov  
Tissue Procurement: APOC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Sequencing by: Inyele Genomics, Inc.  
clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM10293 row: 1 column: 20  
High quality sequence stop: 633.  
Location/Qualifiers  
Source:  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

Search completed: January 19, 2003, 05:06:48  
Job time : 1912.46 secs

---





```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_72"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/vector="pCMV-Sport6, Site:1; NotI;
Site:2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT      344 a      211 c      197 g      291 t
ORIGIN
Query Match      28.8%; Score 910.2; DB 13; Length 1023;
Best Local Similarity 98.4%; Pred. No. 1.6e-167;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

CY 1749 CAAAGTCAGAAATAGACCTTAAATCAAAATCAATTTACAGCATACAGAGATCTCC 1808
DB 1 CAAAGTCAGAAATAGACCTTAAATCAAAATCAATTTACAGCATACAGAGATCTCC 60
CY 1809 TCAAGTCAGAGACATTCAGAGTGTGTGATCAAGAGAGAGAGATGATGATCTAT 1868
DB 61 TCAAGTCAGAGACATTCAGAGTGTGTGATCAAGAGAGAGAGATGATGATCTAT 120
CY 1869 TATAGTATCTATTAAGTCGGTCTTCATCCAGAGATCAAAACAAATATTTCCAGTATGC 1928
DB 121 TATAGTATCTATTAAGTGGTCTTCATCCAGAGATCAAAACAAATATTTCCAGTATGC 180
CY 1929 TCCACAGATCCCTTACCTTGTCCACCAATATCTTACATCTCTCCAGAGCCCTTACAAGT 1988
DB 181 TCCACAGATCCCTTACCTTGTCCACCAATATCTTACATCTCTCCAGAGCCCTTACAAGT 240
CY 1989 TGTATCTTACCTTATGATGATTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 2048
DB 241 TGTATCTTACCTTATGATGATTTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 300
CY 2049 ATATAAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 2108
DB 301 ATATAAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 360
CY 2109 ATATCAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 2168
DB 361 ATATCAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 420
CY 2169 ATATCAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 2228
DB 421 ATATCAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 480
CY 2229 CAAAGATTTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2288
DB 481 CAAAGATTTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
CY 2289 ATATCAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 2348
DB 541 ATATCAATTTGAGAGTCTCCACACACCAAAATGACTCCAGATCAAGATCTT 600
CY 2449 AAGATCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2408
DB 601 AAGATCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
CY 2409 CCAAGTCAGAGATTCAGAGTGTGTGATCAAGAGAGAGAGATGATGATGATGATGAT 2468
DB 661 CCAAGTCAGAGATTCAGAGTGTGTGATCAAGAGAGAGAGATGATGATGATGATGAT 720
CY 2469 TAACCTTCCAGATTCAGAGTGTGTGATCAAGAGAGAGAGATGATGATGATGATGAT 2528
DB 721 TAACCTTCCAGATTCAGAGTGTGTGATCAAGAGAGAGAGATGATGATGATGATGAT 780
CY 2529 AAGATCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2586
DB 781 AAGATCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

```

```

CY 2587 GTATATTATACAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGAT 2646
DB 841 GTATATTATACAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGAT 900
CY 2647 TGTATTTATTTTAAATTAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 2702
DB 961 TGTATTTATTTTAAATTAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 960
CY 2703 AATGTCAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGATGAT 2762
DB 961 AATGTCAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGATGAT 1020
CY 2763 A 2763
DB 1021 A 1021

RESULT 2
BM561083
DEFINITION      1051 bp. mRNA.
AGENCY          NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550905
5' mRNA sequence.
ACCESSION      BM561083
VERSION        BM561083.1 GI:18806043
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Cladialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 1051)
AUTHORS        NIH-MGC http://mgi.mcl.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        C. Robert Strausberg, Ph.D.
                Email: cqaabs@mail.nih.gov
                Tissue Procurement: ATCC
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (IML)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                plate: LHAM12263 row: m column: 18
                High quality sequence stop: 737.
FEATURES
    source
        1..1051
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="NIH_MGC_67"
            /tissue_type="retinoblastoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: eye; Vector: pCMV-Sport6; Site:1; NotI;
            Site:2; SalI. Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.75 kb. Library constructed by Life
            Technologies."
BASE COUNT      330 a      221 c      197 g      299 t      4 others
ORIGIN
Query Match      28.8%; Score 909.4; DB 13; Length 1051;
Best Local Similarity 98.4%; Pred. No. 1.6e-167;
Matches 948; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

CY 607 AAAGTCAGAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGAT 966
DB 1 AAAGTCAGAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGAT 60
CY 967 TGTATTTATTTTAAATTAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1026
DB 61 TGTATTTATTTTAAATTAAGATTTGATGATGATGATGATGATGATGATGATGATGAT 120
CY 1027 AATTTTACAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGATGAT 1086
DB 1027 AATTTTACAGATTCAGAGTGTGTGATCAAGAGAGAGATGATGATGATGATGATGAT 1086

```







```

1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240 3241 3242 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254 3255 3256 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268 3269 3270 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282 3283 3284 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296 3297 3298 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310 3311 3312 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324 3325 3326 3327 3328 3329 3330 3331 3332 3333 3334 3335 3336 3337 3338 3339 3340 3341 3342 3343 3344 3345 3346 3347 3348 3349 3350 3351 3352 3353 3354 3355 3356 3357 3358 3359 3360 3361 3362 3363 3364 3365 3366 3367 3368 3369 3370 3371 3372 3373 3374 3375 3376 3377 3378 3379 3380 3381 3382 3383 3384 3385 3386 3387 3388 3389 3390 3391 3392 3393 3394 3395 3396 3397 3398 3399 3400 3401 3402 3403 3404 3405 3406 3407 3408 3409 3410 3411 3412 3413 3414 3415 3416 3417 3418 3419 3420 3421 3422 3423 3424 3425 3426 3427 3428 3429 3430 3431 3432 3433 3434 3435 3436 3437 3438 3439 3440 3441 3442 3443 3444 3445 3446 3447 3448 3449 3450 3451 3452 3453 3454 3455 3456 3457 3458 3459 3460 3461 3462 3463 3464 3465 3466 3467 3468 3469 3470 3471 3472 3473 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483 3484 3485 3486 3487 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500 3501 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514 3515 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528 3529 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542 3543 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556 3557 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570 3571 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584 3585 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598 3599 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612 3613 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626 3627 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640 3641 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654 3655 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668 3669 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682 3683 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696 3697 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710 3711 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724 3725 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738 3739 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752 3753 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766 3767 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780 3781 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794 3795 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808 3809 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822 3823 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836 3837 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850 3851 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864 3865 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878 3879 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892 3893 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906 3907 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920 3921 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934 3935 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948 3949 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962 3963 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976 3977 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990 3991 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004 4005 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018 4019 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032 4033 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046 4047 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060 4061 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074 4075 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088 4089 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102 4103 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116 4117 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130 4131 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144 4145 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158 4159 4160 4161 4162 
```



























```

RESULT 6
BM450031
LOCUS       1023 bp             11-DEC-91
DEFINITION  RM459031
ACCESSION   AF039344
VERSION     1
KEYWORDS    5' UTR, mRNA SEQUENCE.
SOURCE      Homo sapiens cDNA clone IMAGE:528037
ORGANISM    Homo sapiens

```

VERSION  
KEYWORDS  
SOURCE  
EST.  
human  
004 000 11.1 51:184 9407.1

FILE: NATIONAL INSTITUTE OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)  
JOURNAL: Unpublished (1999)

CONTACT: ROBERT STRAUSBERG, PH.D.  
Email: cspbs\_email@nib.gov

TISSUE PROCUREMENT: ALCO/DCLD/DLP  
cDNA Library Preparations. Life Technologies, Inc.

cDNA Library Arrayed by The I M A G E Consortium (M.I.N.I.)  
DNA Sequencing by, Adenocourt Bioscience Corporation

clone distribution. MCC clone distribution information can be found through the EMAGE database (<http://www.ebi.ac.uk/EMBL/EMBL/>).

http://image.dmi.gov  
Plate: LIAM12204 row: 4 column: 22

FEATURES

source  
1. 1923  
"forquissin"-Homo sapiens"

```
/dev/rtc = 00000000
/clock "IMAGE:528037"
```

```

/tissue_type="melanotic melanoma"

```

```

/patchose -mno (Phage resistance)
/notes="organ: skin; Vector: pCMV-SporkE6; site_1: NotI;

```

Average insert size 2 kb. Library constructed by Life

BASK COUNT	334 a	211 c	197 g	281 t
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

Prod. No.:	1.6e-99	Length:	1023
	1.11 00		0.4
	1.11 00		0.4

Percent Similarity:	100.00%
Conservative:	0

Query Match:	26.81%	Indels:	0
			0

15-00-074-45A-4E (1-707) x DMAE000 (1-1000)

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

[illegible]

60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99

```
574 11eValEbeTurAepSerValEbeMoC1aAraoOutVerThAco11ofow1eTurA1c 541
```

[illegible]

643 Schumfara Proprietate sau Serviciu | 2019-2020 | 2020-2021 | 2021-2022

db 182 TCCATAGGCTCCCACTGGGATTAATACCTACATTCATGAGAGGCTTACCAATTT 241

Ov 662 ProSerSerProLeuAlaIleLeuProC[MCTVAsu]LeuSerProIleuSerPro 681

Db 242 CCTACGTTTACCCCTTACGGATTCTCTGGAGGGAACACTATAATTACCCCTCAAGTACCA 301

[illegible]

Downloaded from <http://ajph.org/> on November 10, 2014





Percent Similarity: 84.41% Conservative: 7  
 Best Local Similarity: 81.75% Mismatches: 48  
 Query Match: 24.93% Indels: 3  
 Gaps: 0

US-09-026-459A-45 (1-797) x AL564757 (1-792)

QY 111 AlaValIleProIleAsnGlySerProArqThrProArqThrGlyGlnAsnArqSerAla 130  
 DB 785 GCTTTATATGATTAATGCTTCAGTGGACACAGAGGAGAGTCAAGCAAGAGTGA 726

QY 131 ArqIleAlaLysIleGlnAsnAspIleArqIleIleGlnValIleGlySerGlnHis 150  
 DB 725 GAGWAGAACAAAYTAGAAATGATACAGCAATATTCAGAGTCTCTGTAACAAAT 666

QY 151 GlnGlyAsnIleAspIleValLysAsnValTyrPheLysAsnPheIleProPheMetAsn 170  
 DB 665 CAATGTAATATACWAGAGGCAAAAAGTWTTCAAAAATTTATACGTTTTAAGAT 696

QY 171 SerLeuGlyLeuValThrSerAsnGlyLeuProGlnValGlnAsnLeuSerLysArqTyr 190  
 DB 605 TTTCTGGACCTTGTMAATCTTWTAGAGCTTCAGAGGTTGAAATTTCTTAAAGATAC 546

QY 191 GlnGlnIleLysLeuLysAsnLysAspLeuAspLysArqIleGlnAsnAspLys 210  
 DB 545 GAAAGAAATTTCTTAAAAAATAAGACATAGATGCAAGATATTTTGGWICATGATAAA 486

QY 211 ThrLeuIleThrAspSerIleAspSerPheGlnThrGlnArqThrProArqLysSerAsn 230  
 DB 445 AATCTTCAGATCATATWTAGACAGCTTCGAAACACAGAGAACACCAAGAAAGTAAAC 426

QY 231 LeuAspIleGlnValAsnValIleProProIleHisThrProValArqThrValMetAsnThr 250  
 DB 425 CTCTWGGAGAGGAGGAAAGTAACTTCAGATATCTGCAATATCTAGTACATTTTGGACAT 466

QY 251 ThrGlnIleLeuMetMetIleLeuAsnSerValAspAspGlnProSerGlnAsnLeuIle 270  
 DB 365 WTCACAAATATATGAGTCTTTAAATTCAGCAAGCTCTCAACCTTCACAAATTTTGTTT 306

QY 271 SerTyrPheAsnGlyThrValAsnProLysIleSerIleGlyArqValLysAsp 290  
 DB 405 TTYWTTTTTACGATCTGACAGTCAATCAAAAGCAAGTATATCAAAAGAGTGAAGGAT 246

QY 291 ThrGlyTyrIleProGlyLeuLysProAlaLysAlaValGlyGlnGlyCysValGlnIle 310  
 DB 245 WTKGALAWTYTTTAAAGAGAAATTTGTAAAGCTGTGGAGACAGGTTGTGTGTCAAAAT 186

QY 311 GlySerIleArqTyrLysLeuGlyValArqLeuTyrTyrArqValMetGlnSerMetL 330  
 DB 185 GATATACACGATACAAAGCTGGTCTCTTCTGCTTGTATACCAAGTAAAGCAATCAATGC 126

QY 331 GlnLysSerGlnGlnGlnArqLeuSerIleGlnAsnPheSerLysLeuLeuAsnAspAsn 349  
 DB 125 TTTATATACAGAGAGAT 33

QY 351 ThrPheHisMetSerLeuLeuAlaCysAlaLeuGlnValValMetAlaThrTyrSerArg 369  
 DB 65 VTTTCTATATGCTTTATATGAGGAGAGGCTTCTTCAATTTTAAATATATATATATATAT 5

QY 370 Ser 370  
 DB 5 ACT 3

RESULT 10  
 BE082846/c  
 LOCUS BE082846  
 DEFINITION BE02-RE08462-219288-014-006-RE0842 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BE082846  
 VERSION BE082846.1 GI:8474151  
 KEYWORDS EST  
 SOURCE Human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 694)  
 Dias Neto, E., Garcia Correa, K., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bondin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Balda, G.S., Simpson, D.H.,  
 Branstetter, A., de Oliveira, P.S., Huetter, P., Jomene, C.V., O'Hare,  
 M.J., Soares, P., Brentani, R.R., Reis, L.F., de Souza, S.J., and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF-expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 4491-4496 (2000)  
 20202664  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/submit2.pl?l=1&l2=K2-ET0642-210-  
 203-914-09647-2-2000-02-23&l=1)  
 Seq primer: puc18 forward  
 High quality sequence start: 58  
 High quality sequence stop: 669.  
 Location/Qualifiers  
 1..694  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="H10042"  
 /dev\_stage="Adult"  
 /note="Organ: Breast; Vector: puc18; Site: 1; Small: Site 2;  
 Small: A mini-library was made by cloning products derived  
 from ORFESTS PCR (U.S. Letters Patent application No. 196  
 7716 - Ludwig Institute for Cancer Research) profiles  
 into the puc18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."  
 BASE COUNT 211 a 129 c 127 g 226 t 1 others  
 ORIGIN

Alignment Scores:  
 Fred. No.: 9.9e-92 Length: 694  
 Score: 1030.00 Matches: 287  
 Percent Similarity: 96.33% Conservative: 3  
 Best Local Similarity: 94.95% Mismatches: 7  
 Query Match: 24.86% Indels: 2  
 DB: 10 Gaps: 0

US-09-026-459A-45 (1-797) x BE082846 (1-694)

QY 275 AsnCysThrValAsnProLysGlnSerIleGlnLysArqValLysAspIleTyrIle 294  
 DB 331 TTTTATATATGCTTTATATGAGGAGAGGCTTCTTCAATTTTAAATATATATATATATAT 33

DB 676 CACGCGACAGGATG-CGAAAGAGAGATATATCTACAGAGATGAGAGATATATATATAT 618

QY 295 PheLysGlnLysPheAlaLysAlaValGlyGlnGlyCysValGlnIleLysSerGlnAla 414  
 DB 617 TTTAAAGAGAAATTTGCTAGAGTGTGGGACATGCTGTGTCGAAATTTGCAATATGATCA 558

QY 315 TyrLysLeuGlyValArqLeuTyrTyrArqValMetLeuSerMetLysArgGlnIle 434  
 DB 557 TCAAACTTGGAGTGGCTTGTATACGAGATATGATGATGATGATGATGATGATGATGAT 498

QY 336 GTCATGAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554  
 DB 497 GAACGATATATGATTCACAAATTTTACCAAACTTCTGCAAACTTCTGCAAACTTCTGCA 488

QY 355 LeuLeuAlaCysAlaLeuGlnValValMetAlaThrTyrSerArgSerThrSerGlnAsn 474  
 DB 437 TTTATGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479

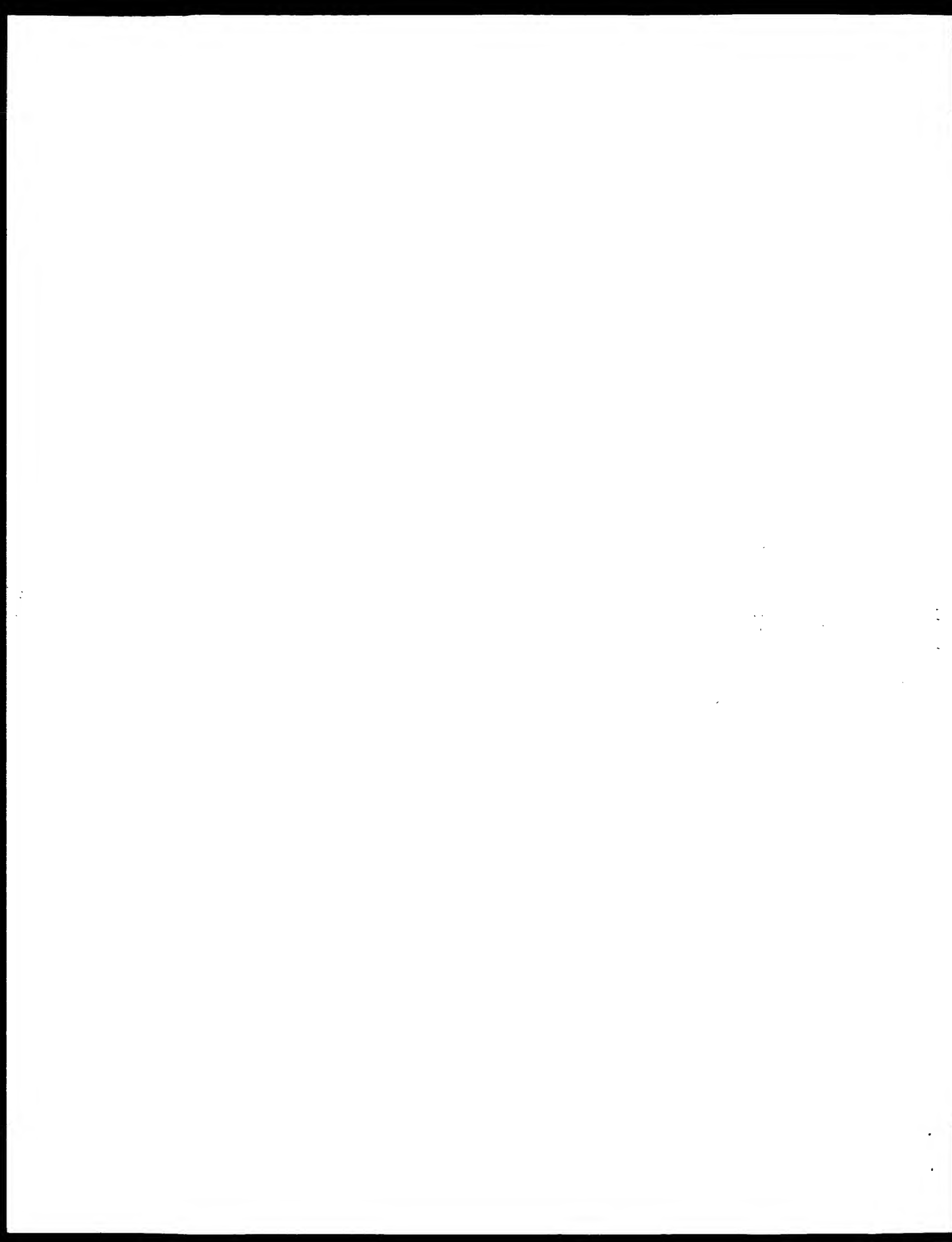














unpublished (1999).  
Contact: Robert Strausberg, Ph.D.  
Email: [csgrbs@mail.nih.gov](mailto:csgrbs@mail.nih.gov)  
Website: <http://www.genome.gov>  
Funding: APOC  
Library preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLUM)  
DNA Sequencing by: Accelerant Bioscience Corporation  
Clone distribution: MAC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILLUM at:

FEATURES	SOURCE
Location/Qualifiers	
1. 1051	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMATH:5550905"	
/clone_lib="NIH_MGC_67"	
/tissue_type="retinoblastoma"	
/lab_host="pBIL10B (phage-resistant)"	
/vector="pUC19"	
Site 2. SalI, Cloned unidirectionally.	
Average insert size 1.75 kb.	
Library constructed by J. C. Miller, Chigo dr.	

BASE COUNT	430 a	321 c	197 f	209 g	4 others
ATTN					
ALIGNMENT SCORES:					
Prod. No.:	1,90	159		length:	1951
Score:	1551.00			Matches:	312
Percent Similarity:	93.26%			Conservative:	6
Best local Similarity:	91.50%			Mismatches:	18
Query Match:	35.39%			Indels:	5
DB:	13			Gaps:	1

435-439	435	439A-439	(1-851) x BM561033 (1-1051)
QY	436	AlaValGlyGlnGlyCysValGluIleLeuGlySerGlnArgTyrLysLeuGlyValArgLeu	375
DB	4	QVPLGSGACACGGTGGTGGTAAATTCGACACACACGACAAACAAATTCGACGTTGGTTC	63
QY	436	TyrTyrArgValMetGluSerMetLeuLysSerGlyGlnIleAlaLeuSerThrLeuAsn	495
DB	64	TATTATCGAGTAATGSAATCCATGCTTAAATTCAGAAAGAAACCATTTATTCATTCAAAAT	123
QY	436	PheSerLysGluLeuAsnAspAsnIlePheHisMetSerLeuAlaLysAlaIleuGln	415
DB	124	TTTAGCAAACTTCGAAGACAAATTCTTATATGTTTATTATGCTGTCCTGAG	183
QY	416	ValValMetAlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeu	435
DB	184	GGTGGAACTGACATATACCAACAACTATCTCAATCTGATTCGCAACACAGATTG	243
QY	436	SerPheProTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIle	455
DB	244	TCCTTCGACAGSATTCTCAATGTCCTTAATTTAAAGATTTTATATTTTCAAAATGATC	303
QY	456	CysSerPheLeuLysAlaTrpLeuLysSerGlnThrAlaGluMetIleLysHisLeuGlyArg	475
DB	304	GAAATTTTATCAAAAGACAAAGACAAATTCACAAAGAAATGATTAATTTAGAACGA	363
QY	476	CysGlnHisArgIleMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeu	495
DB	364	TCGGAATATCAATCATGCAAAAGGCTGATATGGCTTCGACATTCACCTTAATTTGATCT	423
QY	496	IleLeuLysSerLysAspArgGlnIleProThrAspHisLeuGlnSerAlaGlyProLeu	515
DB	424	ATTAAACAATCAAGGACGCAAGACCAATTCATACCTTGAAATTCCTTATGCTCTT	483
QY	516	AsnLeuProLeuGlnAspAspHisThrAlaAlaAspMetTyrIleuSerTyrValArgSer	535
DB	484	AATCTCTCTCAACATAATATATATAGAAATATGATTTTTCCTGTTAAATGAT	543

QY	536	ProLysLysLysGlySerThrThrArgValAspSerThrAlaAsnAlaGlnThrGlnAla	555
DB	544	CCAAACAAAAAAGGTCACACTACGGGTGAAAAATCTACTCTAAATGCACAGACACAAAGCA	604
QY	556	MetSerAlaPheGlnThrGlnLeuLysProLeuLysSerThrSerLeuSerLeuPheLysLys	575
DB	604	AACTATATGCTTCTGAGAGGAGGAAATGAAATCTAACTCTTCTACTGTTTATTATAA	664
QY	576	LysValTyrTyrGlnGlnAlaTyrLeuArgGlnAsnThrLeuGlyGlnAlaGlnLeuSerGln	595
DB	664	AAATCTGATCGCTACCGCTAATCCGGGCTAAATACAGCTTCTGAAACCCCTCTGCTCAG	724
QY	596	HisProGlnLeuGlnGlnHisIleIleThrThrLeuPheGlnHisThrLeuGlnAsnGluTyr	615
DB	724	CACCCACAAATTAACAACATATATCATTCGGACCGCTTTTCACGACACCGTCCAGAAATCAGAT	784
QY	616	GlnLeuMetArgAspArgHisLeuAspGlnIleCysMetCysMetCysMetCysLysLys	635
DB	784	GAACTATACAGACAGAGAGAGGTAATTTGAAATATGAGTGGTCTATGATATGACATATGC	844
QY	636	LysValLysAspIleAspIleLysPheLysIleIleValIleAlaTyrLysAspLeuPyr	655
DB	844	AAATGCAAAATATATACATCTTAANATTCAAAAACCTCTTGAAAGCAACACAGCATCTTCC	904
QY	655	HisAlaValGlnGlnThrPheGlyAspValGlyPheIleLysGlnGlnGlnTyrAspSer	674
DB	904	TCACTGCTGTCACGACACATTCAAAGATGCTGCTTTCAGATCAAGAGGAGGTATGATGAC	964
QY	675	IleIleValPheTyrAspSerValPheMetGlnArgLeuLysThrAsnIle	691
DB	964	TATATAAATAATGCTATAATCTGAGAGCTTTATATTAAGGAGATATAAAAATCAAAAT	1020
RESULT 2			
AK011246			
LOCUS	AK011246	1364 bp mRNA linear	HFC 19-JAN-2002
DEFINITION		Mus musculus 10 days embryo whole body cDNA, RIKEN full-length	
ACCESSION	AK011246	enriched library clone:360017M22 ref 010813ma.1, full insert	
VERSION	AK011246.1	sequence.	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:57BL/6J) 10 days embryo cDNA to mRNA.		
	clone.lib:PIKEN full-length enriched mouse cDNA library		
	clone:360017M22		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;		
REFERENCE			
AUTHORS	Garninci, P., and Hayashizaki, Y.		
TITLE	High efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE			
AUTHORS	Garninci, P., Shibata, Y., Hayatsu, N., Suahara, Y., Shibata, K.,		
	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
	Normalization and subtraction of cap-trapper selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	20499374		
MEDLINE	11042159		
PUBMED			
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
	Konno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M.,		
	Yamamoto, N., Ishii, Y., Nakamura, S., Hayama, M., Nishino, I., Harada, A.,		
	Suzuki, T., Matsumoto, H., Sakaguchi, S., Ikegami, T., Fushiyoshi, K.,		
	Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,		
	Yonekura, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawaji, J.,		
	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
	RIKEN integrated sequence analysis (RISA) system--384-format		
	sequencing pipeline with 384 multicapillary sequencer		
	Genome Res. 10 (11), 1757-1771 (2000)		
JOURNAL	20530913		
MEDLINE			



```

13 133 GATTAAAGCAAAATATTATACATGAAAGATACATCTCTTAATCTATTCAGGTAAATGTTG 792
14 134 CysValLeuAspTyrPheTleuLysLeuSerProPheMetLeuLysCysProTyrLys 163
15 135 TGTGATGAGTATTTTATTAAGTCTCTACCTGTGACATACAGAGAGGATACAAA 852
16 136 ThrAlaValLeuProTleuAsnLysSerProArgThrProArgCysCysAsnArgSer 183
17 137 ACGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 912
18 138 AlaArgTleAlaLysGlnLeuLysAsnThrArgTleLleGlnValLeuLysLysGln 203
19 139 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
20 140 CysLeuLysAsnLysAsnValTyrPheLysAsnPheLleProPheMet 223
21 141 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
22 142 AsnSerLeuGlyLysValThrSerAsnGlyLeuProLysValLysAsnLeuSerLysArg 243
23 143 AATTCATGCGAATGCGAATGCGAATGCGAATGCGAATGCGAATGCGAATGCGAATGCG 1092
24 144 TTTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 263
25 145 TATCAAGAGATTTATCTTAAACAAAGATTTATGATGAAAGCTGTTTGTGATATGAT 1152
26 146 CysLeuLysSerProPheCysSerPheCysThrGlnAlaPheProArgLysSer 283
27 147 AAAATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
28 148 AsnLeuAspGlnValValAsnValLysProProHisThrProValArgThrValMetAsn 303
29 149 AAGCCATGATCAACAGCCATGATCAACAGCCATGATCAACAGCCATGATCAACAGCCAT 1272
30 150 ThrLleGlnLeuMetMetLleLeuAsnSerAlaSerAspGlnProSerGlnAsnLeu 323
31 151 ACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1332
32 152 IleSerLysPheAsnAsnGlySerThrValAsn 343
33 153 ATTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1392

```

```

RESULT 3
HQ220275
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
COMMENT

```

```

881 bp cDNA linear EST 02 MAY 2002
5' mRNA sequence.
cDNA library preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: M3 clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13287 row: c column: 03
High quality sequence stop: 690.
Location/Qualifiers
1..881
/oranism="Homo sapiens"

```

```

/db_xref="taxon:4606"
/clone IMAGE:6044522
/clone_lib="NIH_MCC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Orqan: testis; Vector: pCMV-SPORT6; Site: 1: Not 1;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MCC library."
HASH COUNT 313 a 149 c 155 g 264 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 7,530-140 Length: 881
Score: 1448.00 Matches: 288
Percent Similarity: 98.30% Conservative: 1
Best Local Similarity: 97.96% Mismatches: 3
Query Match: 33.04% Indels: 2
DB: 24 Gaps: 0

```

US 09-026-459a-39 (3 851) x HQ220275 (1-881)

```

QY 131 MetGluAspAspLeuValLysSerPheGlnLeuMetLeuLysValLeuAspTyrPheLle 150
DB 3 ATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 62
QY 151 CysLeuSerProPheMetLeuLysSerProPheMetLeuLysSerProPheMetLeuLys 170
DB 63 AAAATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
QY 171 GlySerProArgThrProArgArgGlyGlnAsnArgSerAlaArgLysLeuValLeu 190
DB 123 GGTTCACCTGCAACAGCCATGATCAACAGCCATGATCAACAGCCATGATCAACAGCCAT 182
QY 191 GluAsnAspThrArgLleLleGlnValLeuLysLysGlnHisGlnCysAsnLysAspGln 210
DB 183 GAAATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242
QY 211 ValLysAsnValTyrPheLysAsnPheLleProPheMetAsnSerLeuGlyLeuValThr 230
DB 243 GTGAAAATGATTTATTTCAAAAATTTATGATTTATGATTTATGATTTATGATTTATGAT 302
QY 231 SerAsnGlyLeuProGlnValGlnAsnLeuSerLysArgTyrGlnGlnLleTyrLysGln 250
DB 303 TCTAATGCACTTCCACAGCTTCAAAAATCTTCTAATAACATACAGAGAAAATATCTTAAA 362
QY 251 AsnLysAspLeuAspAlaArgLeuPheLeuAspGlnLysAspLysThrLeuGlnThrAspSer 270
DB 363 AATGAAATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
QY 271 IleAspSerPheGlnThrGlnArgThrProArgGlySerAsnLysAspGlnGlnValAsn 290
DB 423 ATAGAAGTCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 291 ValLysProProHisThrProValArgThrValMetAsnThrLleGlnGlnLeuMetMet 310
DB 483 GAAATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
QY 311 IleLeuAsnSerAlaSerAspGlnProSerGlnAsnLeuLysSerTyrPheAsnAspGln 330
DB 543 ATTTTAAATTCACCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
QY 331 ThrValAsnProLysGlnSerLleLysLysArgValLysAspLleGlyTyrLlePheLys 350
DB 603 AATGAAATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
QY 351 GlyLysPheValLysValGlnGlyLysValGlnGlySerGlnArgThrLys 370
DB 663 CACAAATTTCTTAAAGCTGTGGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 722
QY 371 LeuGlyValArgLeuTyrTyrArgValMetGlnSerMetLeuLysSerGlnGlnLeuArg 390

```

224 CTTGGAGTTGGCTTGATTAACGAGTAAGAAATGCGATGTTAAATGCAAGAGAGACGGA 782  
 391 leuSerThrGlnAsnGlySerThrLeuLeuAspAsnGlyGlyHisMetSerLeuL 410  
 784 TATGCAATGAAAAATTAAGTAAATTCGGAATGACACACATTTTTCATATGCTTTAT 842  
 410 euAlaCysAlaLeuGluValValMetAlaThrTyrSer 422  
 844 TGGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880  
 RESULT 4  
 00475603 1001 bp mRNA linear EST 05-FEB-2002  
 DEFINITION A58600007 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5575683  
 5', mRNA sequence.  
 ACCESSION 00475603  
 VERSION 00475603.1 GI:18524645  
 KEYWORDS EST  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1001)  
 AUTHORS NIH MGC project, March 2001  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rob@nhi.nih.gov  
 Issue Procurement: A100  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution by: Aducent Bioscience Corporation  
 Found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 plate: L1A1246 row: 1 column: 04  
 High quality sequence start: 17  
 High quality sequence stop: 734  
 location/qualifiers  
 1. 1001  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5575683"  
 /clone\_lib="NIH\_MGC\_92"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="BHK21 (phage-resistant)"  
 /note="organ: testis; vector: pCMV-Sport6; Site: 1; Not:  
 Site 2; Salt: cloned unidirectionally; oligo-dT primed  
 Average insert size 2.5 kb. Library enriched for  
 full length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC library.  
 BASE COUNT 449 a 176 c 186 g 288 t  
 ORIGIN  
 Alignment Scores:  
 Seq. No. Length: 1001  
 Scores: 1449.00 Matches: 281  
 Percent Similarity: 94.87% Conservativity: 16  
 Best Local Similarity: 90.65% Mismatches: 10  
 Query Match: 90.65% Indels: 9  
 Gaps: 4  
 US 09 026 459A 39 (1 851) x 00475603 (1 1001)  
 07 58 leuLysGluThrAspHisSerThrLysValAspAsnAlaMetSerArgleuLeuLysLys 77  
 14 68 GTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 97  
 27 78 TTTAspValLeuGluAlaGlySerLysLeuGluArgThrCysGlnLeuLeuLysLys 97  
 46 98 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 157  
 97 97 GThreGluProSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 117

158 GACACACACGCACTTCGATATCTACTGAAATAAAATTCAGATGAGG 217  
 117 rtrpIleThrPheLeuLeuAlaLysGlyGluValLeuGluMetGluAspLeuVal 147  
 218 TTGGATCAATTTTTTATTATAGGTAAAGGGGAAATATTAAATGGAAATGATTTTGGAT 277  
 137 eSerPheGlnLeuMetLeuLysValLeuAspTyrPheIleLysLeuSerProGluMetLe 167  
 278 TTTCAATTCAGTTAAAGCTATGCTGCTTACATCTTAAATTTTAAATTTTAAATTTTAA 347  
 157 leuLysGluProTyrLysThrAlaValIleProIleAsnGlySerProArgThrProArg 177  
 348 GGTCAAGAAACCATATAAAGAGGTGTATATACCATTAATGTTTCACTCGAAATCCAG 497  
 177 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 197  
 398 GCGAGTCAGAACAG 457  
 197 GCGuValLeuGlySerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 217  
 458 TCAAGTTCCTCTCTAAAGAACATGAATGATGATGATGATGATGATGATGATGATGATGAT 517  
 217 sAsuThrIleThrIleMetAsuThrLeuValIleThrSerAsuThrLeuThrIleVal 247  
 518 AAATTTTATAGCTTTTATGAAATCTCTTGGATTTGTAACATCAATGACATTCGAAAGAT 577  
 237 TCTuAsuLeuSerLysAlaTyrGluGluIleLysLysLysLysLysLysLysLysLysLys 267  
 578 TCAAAATCTTTCTAAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637  
 257 qLeuPheLeuAspHisLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 277  
 638 ATATTTTGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 697  
 277 uAlaThrProAlaLysSerAsuLeuAspGluValAsuValIleProHisThrPro 297  
 698 GAGAAATCTGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 757  
 297 GValArgThrValMetAsuThrIleGlnGlnLeuMetIleLeuAsuSerAlaSerAs 317  
 758 AGTTAGAGACTGTTATGAACTATGCAACATTAATGATGATGATGATGATGATGATGATGAT 817  
 317 pGlnProSerGluAsnLeuIleSerTyrPheAsnAsuLysThrValAsuProLysGlu 346  
 818 TCAACTTTTCAAAAAATGCTATTTTCTAATTTTAAACACATGCAATGCAATGCAATGCA 877  
 346 uSerIleLeuLysArg---ValLysAspIleLeuTyrIle---PheLysLysLys 392  
 878 AAGCTATACCGGAG 947  
 353 ---PheAlaLysAlaValGly 35H  
 948 TTCTGGCTAAGAGCTGGG 957  
 RESULT 5  
 00475603 727 bp mRNA linear EST 14 AUG 2001  
 DEFINITION DKEP41300816\_1 43 (synthetic clone) Homo sapiens cDNA clone  
 DKEP41300816\_5', mRNA sequence.  
 ACCESSION 00475603  
 VERSION 00475603.1 GI:15160502  
 KEYWORDS EST  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 727)  
 AUTHORS Postek, A., Wellenreuther, K., Moses, H.W., Well, B., and Wichmann, S.  
 TITLE EST (Postek, A., Wellenreuther, K., Moses, H.W., Well, B., and Wichmann, S.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Postek, A.







```

CY 674 AspSerIleValPheTyrAsnSerValPheMetGlnArqLeuLysThrAsnIleLeu 692
DB 675 SAHICATATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 667
CY 693 GlnTyrAlaSerThrArgProPro---ThrLeuSerProIleProHISile-ProArgSe 711
DB 694 CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 707
CY 711 ProlLys---PheProSerProLeu---ArgIlePro---GlyGlyAsn----- 726
DB 712 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 707
CY 727 IleTyrIleSerProLeuLys 733
DB 728 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727
CY 748 TATATATATATATATATATATATATATATATATATATATATATATATATAT 809
DB 749 TATATATATATATATATATATATATATATATATATATATATATATATATAT 809

RESULT 8
BM450041 1023 bp mRNA linear EST 05-FEB-2002
ACEM-001 6493544 NIH-MGC-72 Homo sapiens cDNA clone IMAGE:5528037
5' mRNA sequence.
ACCESSION BM450041
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1023)
AUTHOR Nishimura, H., et al.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstraus@nhi.nih.gov
Tissue Procurement: ATCC/DCMP/DBP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI)
DNA Sequencing by: Amersham Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNI at:
http://image.llni.gov
Plasmid: LLMI2204 row 3 column: 22
High quality sequence stop: 643.
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="BM450041"
/issue_type="NIH-MGC-72"
/lab_host="DH10B (phage-resistant)"
/site="Orig. skin. Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE 1 UNT 44 a 211 c 197 q 281 t
ORIGIN

Alignment Scores:
Prod No.: 118e-104 Length: 1023
Score: 111.00 Matches: 216
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.35% Indels: 0
DB: 13 Gaps: 0

US-09-026-459a-49 (1-851) x BM450041 (1-1023)
CY 695 LysValLysAsnIleAspLeuLysPheTyrIleValThrAlaTyrLysAspLeu 655
DB 696 AAAGTGAAGAAATATACACCTTAAATTCATGATGATGATGATGATGATGATGAT 641
CY 696 HisAlaValGlnGluThrPheLysArgValLeuLysLysGluGluThrAspSerIle 675

```

```

DB 62 CATGCTGTCAGAGACATCAAAAGTGTGTTGATCAAGAGAGAGATGATGAT 141
CY 676 IleValPheTyrAsnSerValPheMetGlnArqLeuLysThrAsnIleLeuLysPhe 695
DB 122 ATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 141
CY 696 SerThrArgProProThrLeuSerProIleProHISile-ProArgSerProLysPhe 715
DB 182 TCCACGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
CY 716 ProSerSerProProArgIlePheGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 745
DB 242 CCTAGTTCACGCTTACGGATTCCTGGAGGAGACATCATATATTCACCTGAGAGATCCA 401
CY 736 TrrLysIleSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 765
DB 302 TATATAATTTTCAAGAGTCTGCTCAACACCAACCAAAATGACTCCAGATGATCAAGATGAT 461
CY 756 ValSerIleGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 775
DB 462 GATCAATTCGCTGATCATTCGCGACTTCGAGAGATTCAGAGATTCAGAGATTCAG 441
CY 776 CysAsnSerAspArgValLeuLysArgSerAlaGluGlySerAsnProProlLysProLeu 795
DB 422 TGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
CY 796 LysLysLeuArgPheAspIleGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 815
DB 482 AAAAATATGCTTTGATATGAGAGATCAATGAGAGATCAATGAGAGATCAATGAG 541
CY 816 GlyGlySerLysPheGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 835
DB 542 GGCACATCCAAATTTTCAGTCAAAATTCGCAAGAAATGATTCATTCACATGCAACAGAA 601
CY 836 LysGlnLysMetAsnAspSerMetAspThrSerAsnLysGluGluLys 851
DB 602 AAGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649

RESULT 9
AA763411 914 bp mRNA linear EST 27-JAN-1998
v54a04 r1 Soares_mammary_gland_NMLK6 Mus musculus cDNA clone
IMAGE:1247598 5' similar to gb:M15400 RETINOBLASTOMA-ASSOCIATED
PROTEIN (HUMAN); gb:M26391 Mouse retinoblastoma susceptibility
protein (MOUSE);, mRNA sequence.
ACCESSION AA763411
VERSION AA763411
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 914)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubouche, J.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, P., Underwood, K., Moore, H.,
Theising, B., Wylie, T., Lennon, G., Soares, P., Wilson, R. and
Waterston, R.
The WashU-HUMI Mouse EST Project
Unpublished (1996)
Contact: Marra, M/Mouse EST Project
WASHU-HUMI Mouse EST Project
Washington University School of MedicineP
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through HUMI; contact the
IMAGE Consortium (info@image.llni.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 422.

```

```

FEATURES
  source          location/qualifiers
  1..914
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /clone_lib="MAGE:1247598"
  /clone_lib="Sources_mammary_gland_NMLMC"
  /sex="female (lactating)"
  /tissue_type="mammary gland"
  /lab_host="DRI106"
  /note="Vector: pT7T-G. Pac (Pharmacia) with a modified
polylinker: 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Ronaldo."
BASE COUNT      257 a 237 g 261 t 1 c others
ORIGIN
Alignment Scores:
  Pred. No.:      4,656/99      Length:      914
  Score:          1956.50      Matches:      244
  Percent Similarity: 82.14%      Conservative: 19
  Best local Similarity: 75.97%      Mismatches: 54
  Query Match:    24.11%      Indels:      6
  Gaps:           2
  DB:
US 09-026-459a-39 (1-851) x AA763411 (1-914)
27 259 phtcousAspHisAspGlyThrLeuGlnThrAspSerHisAspSerGluThrGlnArg 278
Db 914 TTTTGGATGACGAAATAAAGCTTGAGCTGT--TCCTTTAGGGCTTTTAAAGCGGGAG 857
279 ThrProAlaLysSerAsnLeuAspGluGluValAsnValileProProHisThrProVal 298
Db 856 AAGCAAGCAAAACACATCTCTCAAGCGGCAAAAGCGGTACGCGCCACCCCGGT 798
299 ArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGln 318
Db 797 TACGCTTTTITGAAAGCATATACAAATATTTATGTGATTTTAAAGCTGGCAATAGCGG 738
319 ProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProLysGluSerIle 338
Db 747 CACTCAGAAATGCTTT TGGCTCTCTATATTTGGGCGGATGACCAAAAGAAATTCG 679
339 LeuLysArgValLysAspIleGlyTyrIlePheLysGlnLysPheAlaLysAlaValGly 358
Db 678 CTAAGGTCAGGCTTAAAGCAATGCGGACATCTCTTAAACAGCAAGTCTCTA---ACGCTCTCG 623
359 GluGlyGlyValAlaIleIleCysSerIleAcATyGlnLeuGlyValAlaAlaLeuTyrArg 378
Db 622 CAGAGTGGTGAGCATGAGGAGTACAGCGATATAAATTTGAGTCGATGCTATTACGGT 563
379 ValMetGluSerMetLeuLysSerGluGlnThrArgLeuSerIleGlnAspHisSerLys 398
Db 562 GGTATGGAATCATACATTATACATACAGACAGACAGAGTTTGTCATCAGATTTTAGCAG 503
399 LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValMet 418
Db 502 CTCTTAAATGACAAACATCTTTCATATGCTCTTACGCGCTCTGCTTCAAGTCTAATG 443
419 AlaThrTyrSerArgSerThrSerGluAsnLeuAspSerGlyThrAspLeuSerPhePro 438
Db 442 GTATGATATGAGCAAGATGATATTTGATTTGATTTGAGAAAGAAATTTGCTGCGG 365
439 TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPhe 458
Db 382 TCAATTCTCAAGATTAATTTAAAGGCTTTTTCATTTTACAAAGTCATGGAAGTTT 323
459 IleLysAlaGluGlyAsnLeuThrArgGlnMetIleLysHisLeuGluArgCysGluHis 478
Db 322 ATCAAGCTGGCAAGATTAATGCAAGACAAATGATTAAGAAATTTTACAAATCTGACCAT 263

```





This sequence was available from the IMAGE Consortium, for clone  
 orders contact: info@image.llnl.gov  
 Seq primer: -40BP from Gibco  
 High quality sequence stop: 430.

# FEATURES

Location/Qualifiers

1..623

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="HWA:553049"

/clone\_lib="Human Pancreatic Islets"

/issue\_type="Islets of Langerhans"

/lab\_stud="Adult"

/lab\_host="DH10B"

/note="Origin: Pancreas. Vector. ESPORT1 Site 1: Not 1; Site 2: Sal 1; Library constructed using Superscript Plasmid library kit (Life Technologies). cDNA made by oligo-dT priming. Size selected by column fractionation. average insert size 1.08 kb. Primary library. unamplified."

BASE COUNT 234 a 43 c 109 g 186 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 4,090-95 Length: 623

Score: 1016.00 Matches: 201

Percent Similarity: 98.07% Conservations: 2

Best Local Similarity: 97.10% Mismatches: 3

Query Match: 23.19% Indels: 1

DB: 10 Gaps: 0

US-09-026-459a-39 (1-851) x AW584181 (1-623)

QY 104 ThrArqIlelledValLeuGysLyscluhfisclycysAsnIleAspIluVallysAsn 213

DB 3 ACAAATATTAAGTCTCTATANAAGCATAACTAATAATAGATGATGAAAAAT 62

QY 214 ValTyrThelcysAsnPhelcPheMcAcAsn-SerLeuGlyLeuValThrSerAsnGl 233

DB 64 GTTATTTCAAAAATTTTATACGTTTTATGATGTTGCTCTGACCTTATCTAATG 122

QY 235 yLeuPheAlaValAlaLeuSerLysAsqTyGluGluTyLeuLysAsnLysAs 253

DB 123 ACTTCACAGTTCACAAATCTCTAAACCATACGACGAAATTTATCTTAAAAATAACA 182

QY 254 pLeuAsnAlaArqLeuPheLeuAspHisAspLysThrLeuGluThrAspSerIleAspSe 273

DB 184 TATAATATTAATTTTGTGATATATAAATTTCTTCAACTATTTCTATATATATAT 242

QY 273 rPheGluThrGluArqThrProArgLysSerAsnLeuAspGluGluValAsnValIleDr 293

DB 243 TTTTAAACACAGAGACATACGAAAAATATCTTTGATGAGAGATGATGTAATTC 402

QY 294 oProHisThrPheValArqThrValMetAsnThrIleGluGluLeuMetIleLeuAs 313

DB 303 TCCATACACTCCATACGATGATATGACACACTATCCAAATTAATCAATATTTAAA 362

QY 313 userAlaSerAspIleProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAs 333

DB 363 TTCACAGAGATACACCTTCAGAAAATCTGATTTCTGATTTTATAGCAACTGCACGTAA 422

QY 363 rPheGluThrGluArqThrProArgLysSerAsnLeuAspGluGluValAsnValIleDr 393

DB 423 TCTTAAACACAGAGACATACGAAAAATATCTTTGATGAGAGATGATGTAATTC 482

QY 393 eAlaLysAlaValGlyGlnGlyLysValGluIleGlySerGluArgTyLysLeuGlyVa 373

DB 483 TCTTAAACACAGAGACATACGAAAAATATCTTTGATGAGAGATGATGTAATTC 542

QY 474 LArqLeuTyrTrpArqValMetGluSerMetLeuLysSerGluGluGluArqLeuSerIl 393

DB 543 TGGCTTGTATTAGCGATATGAAATCATGCTTAATAATCAGAAAGAAAGATATTGCA 602

QY 493 cGlnAsnPhcSerLysLeu 399

DB 603 TCAAAATTTTACCAACTT 621

## RESULT 14

AA763485/c

LOCUS

DEFINITION

AA763485

871 bp mRNA linear EST 27 JAN-1998

IMAGE:1247665 5' similar to gb:M15403 RPTINOHPLASTOMA ASSOCIATED

PROTEIN (HUMAN), gb:M26392 Mouse retinoblastoma susceptibility

protein (MOUSE);, mRNA sequence.

AA763485

KEYWORDS

EST.

AA763485.1 GI:2813232

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;

(bases 1 to 871)

Marra,M., Hillier,L., Allen,M., Howles,M., Dietrich,N., Dubouque,L.,

Geisel,S., Kucuba,T., Lacy,M., Le,M., Martin,J., Kellis,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU HHMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

tel: 314 286 1800

fax: 314 286 1810

E-mail: mouseest@watson.wustl.edu

This clone is available royalty free through LORIS. Contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

ML:561353

Seq primer: -28ml3 rev2 ET from Amersham

high quality sequence stop. 495.

location/Qualifiers

1..871

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:1247665"

/clone\_lib="Soares\_mammary\_gland\_NMLMG"

/sex="female (lactating)"

/issue\_type="mammary gland"

/lab\_host="DH10B"

/note="Vector. pT73D-Pac (Pharmacia) with a modified

polylinker. 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adapters (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

modified pT73 vector. Library is normalized. Library

was constructed by Rento Soares and M. Fatima Ronaldo."

BASE COUNT 242 a 172 c 204 g 251 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 1,450-94 Length: 871

Score: 1013.00 Matches: 216

Percent Similarity: 80.34% Conservations: 17

Best Local Similarity: 74.48% Mismatches: 53

Query Match: 23.12% Indels: 5

DB: 9 Gaps: 2

US-09-026-459a-39 (1-851) x AA763485 (1-871)

QY 279 ThrProArgLysSerAsn--LeuAspGluGluValAsnValIleProHisThrProV 298

DB 860 AGCTTAAAGTAAAAATAACTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804

QY 298 aArqThrValMetAsnThrIleGlnGlnMetMetIleLeuAsnSerAlaSerAsp 318



```
801 pLLeuGluLysSerAspPheAlaAspGlySerIleHisLeuProGlyGluSerLysPhe 820
|||||
596 TATTCAGGCAATCAGATGAGGAGAGAAATGAGTAAATCCCTCCAGAGAGATCCCAATCT 655
|||||

821 GlnGlnLysLeuAlaGluMetThrSerThrArgThrMetGlnLysGlnLysMetAsn 840
|||
656 TATGCTCAAAATGAGGAGAAATGAGTATTTGAAAAAGAAATGCTAAAGGGAAGATCAAT 715
|||||

841 AspSerMetAspThrSerAsnLysGluGluLys 851
|||
716 GGAAGCTGGTACCTCCACACAGGGAAGAAAG 748
|||||
```

Search completed: January 19, 2003, 05:06:23  
Job time : 1895.76 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd

M nucleic nucleic search, using SW model

Run on: January 16, 2003, 15:20:22 / Search time 3044.87 Seconds  
(without alignments)  
18469.506 Million cell updates/sec

Filter: US 09 026 459A 40

Perfect score: 3461

Sequence: 1 GGGGCAAGCGGGGAAAAC.....AAATGACATATATCACT 3461

Scoring table:

IDENTITY\_NP\*

Gapop 10.0, gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 42408142

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 6%

Maximum Match 100%

Listing first 45 summaries

Database:

EST:\*

1: cm\_estb1a.\*

2: cm\_estb1b.\*

3: cm\_estb1c.\*

4: cm\_estb1d.\*

5: cm\_estb1e.\*

6: cm\_estb1f.\*

7: cm\_estb1g.\*

8: cm\_estb1h.\*

9: cm\_estb1i.\*

10: cm\_estb1j.\*

11: cm\_estb1k.\*

12: cm\_estb1l.\*

13: cm\_estb1m.\*

14: cm\_estb1n.\*

15: cm\_estb1o.\*

16: cm\_estb1p.\*

17: cm\_estb1q.\*

18: cm\_estb1r.\*

19: cm\_estb1s.\*

20: cm\_estb1t.\*

21: cm\_estb1u.\*

22: cm\_estb1v.\*

23: cm\_estb1w.\*

24: cm\_estb1x.\*

25: cm\_estb1y.\*

26: cm\_estb1z.\*

27: cm\_estb1aa.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.2	26.3	1023	13	BM450031 AGENCOURT
2	909.4	26.3	1051	13	BM561083
3	896	24.7	1081	14	BQ439221 AGENCOURT
4	844.8	24.4	881	14	BQ220275 AGENCOURT
5	790	22.9	1344	11	AK011246 Mus muscv
6	786	22.7	1001	13	BM475603 AGENCOURT

7	752.4	21.7	1122	13	BM545881
8	725.4	21.0	727	9	AL597811
9	720.5	20.8	883	14	BQ220275
10	699.6	20.2	823	12	BQ610661
11	698.2	20.2	792	9	AL563757
12	692.2	20.0	747	12	BQ570456
13	691.8	20.0	972	13	BM466194
14	678	19.6	845	13	BM453724
15	660	19.1	672	13	B1868459
16	643.6	18.6	797	10	BE549278
17	642.3	18.6	681	12	BQ616219
18	617	17.8	618	13	BM264009
19	614	17.7	694	10	HE082846
20	605.2	17.5	623	10	AW583181
21	605	17.5	605	9	AL599633
22	594.4	17.2	922	12	BQ253543
23	580	16.8	580	9	AL599105
24	575.2	16.6	914	9	AA763411
25	566.4	16.4	588	10	AW368244
26	558	16.1	871	9	AA763485
27	530.2	15.3	598	20	AW582967
28	530	15.3	549	9	AL692790
29	528	15.3	572	9	AL139941
30	516.8	14.9	597	12	BE931246
31	513.2	14.8	594	10	AV715533
32	512.2	14.8	744	12	BF144857
33	505.2	14.6	746	13	B1151568
34	493.8	14.3	525	9	AA258255
35	486.4	14.1	584	12	BG149050
36	484.4	14.0	493	10	AW502887
37	481.4	13.9	484	10	BE168095
38	480.2	13.9	659	13	B1211116
39	476.6	13.8	2083	13	BM456728
40	474.4	13.7	744	13	B1526982
41	459.2	13.3	497	12	BE931284
42	457	13.2	457	9	AL093215
43	457	13.2	742	9	AL598766
44	454.2	13.1	869	12	HE865749
45	447	12.9	447	9	AL082179

#### ALIGNMENTS

RESULT 1  
BM450031  
LOCUS  
DEFINITION AGENCOURT\_6393544 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5528037  
5', mRNA sequence.

ACCESSION BM450031

VERSION BM450031.1 GI:18499971

KEYWORDS EST.

SOURCE Human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 1023)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cvapbs@mail.nih.gov](mailto:cvapbs@mail.nih.gov)

Tissue Procurement: ATCC/DCTD/UTP

cDNA Library Preparation: Life Technologies, Inc.

Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLNL2204 row 3 column 22

High quality sequence stop: 643

Location/Qualifiers 1..1023

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone "IMAGE:5550905"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/auth="Orqan, skib, Vector, PCMV-Sport6, Site_1, NotI,
Site_2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT      344 a      211 c      197 q      281 t
ORIGIN
Query Match      26.1%      Score 910.2; DH 13; Length 1023;
Best Local Similarity 95.4%; Pred No. 7 to 168;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

QY 2049 CAAGAGCAAGAAATATACAGCTTAAATTCAGAAATTCAGAGGATACAGGATCTCC 2108
DB 1 CAAGAGCAAGAAATATACAGCTTAAATTCAGAAATTCAGAGGATACAGGATCTCC 60
QY 2109 TCATGCTTTTCAGAGAAATTCAGAAATTCAGAGGATACAGGATCTCC 2169
DB 61 TCATGCTTTTCAGAGAAATTCAGAAATTCAGAGGATACAGGATCTCC 120
QY 2169 TATATATCTTATATATCTTATATCTTATATCTTATATCTTATATCTTAT 2228
DB 121 TATATATCTTATATATCTTATATCTTATATCTTATATCTTATATCTTAT 180
QY 2229 TCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 2288
DB 181 TCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 240
QY 2289 TCTAGTATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 2348
DB 241 TCTAGTATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 300
QY 2449 ATATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 2408
DB 401 ATATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 460
QY 2469 ATATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 2468
DB 461 ATATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 420
QY 2469 ATATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 2528
DB 421 ATATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 480
QY 2529 CAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 2588
DB 481 CAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 540
QY 2589 AGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 2648
DB 541 AGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 600
QY 2649 AAATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 2708
DB 601 AAATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 660
QY 2709 CAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 2768
DB 661 CAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 720
QY 2769 TATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 2828
DB 721 TATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 780
QY 2829 AAATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 2888
DB 781 AAATATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 840

```

2887 GTTATTTATACAAAGATGAAATTTCTGTCTTAAATCTCTCCATTTAAAGATGAGTAA 2946  
 DB 841 GTTATTTATACAAAGATGAAATTTCTGTCTTAAATCTCTCCATTTAAAGATGAGTAA 900  
 QY 2947 TTTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3002  
 DB 901 TTTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960  
 QY 3003 AAT 3062  
 DB 961 AAT 1020  
 QY 3063 A 3063  
 DB 1021 A 1021

RESULT 2  
 RM561083  
 LOCUS  
 DEFINITION RM561083 1051 bp mRNA linear EST 20 FEB 2002  
 5' mRNA sequence.  
 ACCESSION RM561083  
 VERSION RM561083.1 GI:18806043  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1051)  
 AUTHORS NIH-MGC http://mgi.mcg.mcg.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Straussberg, Ph.D.  
 Email: rqa@bbs.mcg.mcg.gov  
 Tissue Procurement: ATCC  
 cDNA library preparation: Life Technologies, Inc.  
 cDNA library arrayed by: The I.M.A.G.E. Consortium  
 DNA sequencing by: Agencourt Bioscience Corporation  
 clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 plate: LAM12263 row: n column: 18  
 High quality sequence stop: 737.

FEATURES  
 source  
 1. 1051  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5550905"  
 /clone\_lib="NIH\_MGC\_67"  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Orqan: eye; Vector: PCMV-Sport6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies."

BASE COUNT 330 a 221 c 197 q 299 t 4 others  
 ORIGIN

Query Match 25.1% Score 909.4; DH 13; Length 1051;  
 Best Local Similarity 98.4%; Pred No. 10-167;  
 Matches 948; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1207 AAAAT 1266  
 DB 1 ATAGCT 60  
 QY 1267 TTTTAT 1326  
 DB 61 TTTTAT 1386  
 QY 1327 AATTTAT 1386























```

/db xref: "taxon:9606"
/clone "MAHE:549522"
/clone_id "NIH_MGC_90"
/issue_type "adenocarcinoma, cell line"
/lab host "DH10B (phage-resistant)"
/notes "Organ: Liver; Vector: pCMV-Sport6; Site_1: Not1;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

```

```

BASE COUNT      228 a      124 c      134 g      186 t
CGATGTT

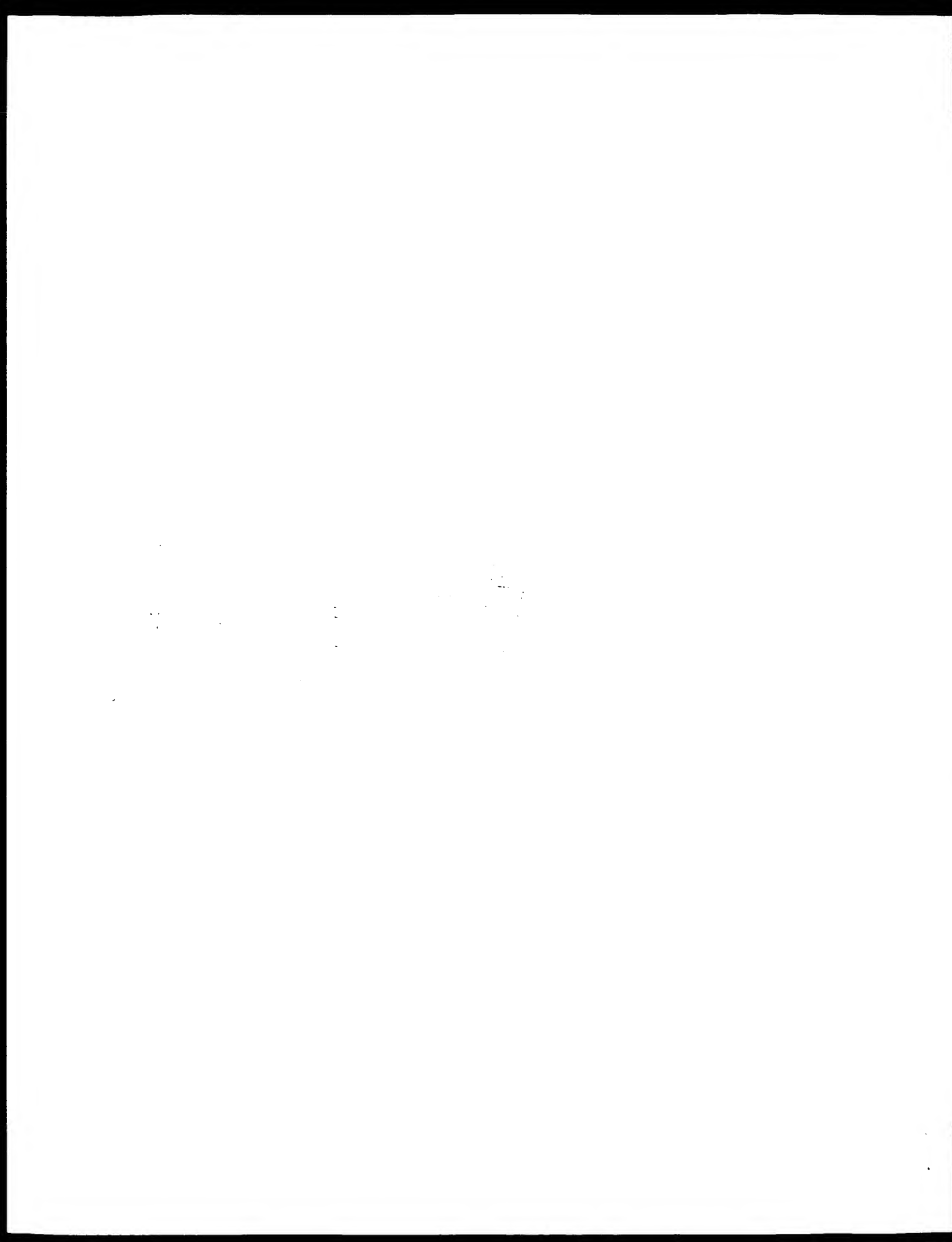
```

```

Query Match      19 18; Score 660; DB 13; Length 672;
Best Local Similarity 99.98; Prod. No. 6,46-119;
Matches 671; Conservative 0; Mismatches 0; Gaps 1;
27 2327 ATATTTATCTGTAAGTGGATATATAATTTGAGAGCTCTGCCAACAACCAACAAA 2386
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 60
27 2387 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2446
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 60
27 61 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 120
27 2447 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2506
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 180
27 2507 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2566
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 240
27 2567 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2626
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 300
27 2627 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2686
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 360
27 2687 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2746
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 420
27 2747 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2806
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 480
27 2807 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2866
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 540
27 2867 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2925
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 600
27 2926 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 2985
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 660
27 2986 TGATCTGAAGATCAAGAAATTAATTAATTAATTTGAGAGCTCTGCCAACAACCA 3044
DB 1 ATATTGACCGCTGAAGAGTGGATATATAATTTGAGAGTCTGCCAACAACCAACAAA 672

```

Search completed: January 18, 2003, 04:09:18  
 Job time: 4064.96 secs





Carninci, P., and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 99279253  
 MEDLINE  
 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 MEDLINE  
 11042159  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishi, K., Kutsunai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishimi, T., Harada, A.,  
 Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Iwaga, K., Iwaga, Y., Iwaga, M., Ohara, E., Watabiki, M.,  
 Yacoda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawaji, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-484-format  
 sequencing pipeline with 484 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20539913  
 MEDLINE  
 11076861  
 4  
 Kawaji, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Aizawa, K., Iwaga, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, S.,  
 Saito, K., Okazaki, Y., Goto, T., Bono, H., Kasukawa, T., Saito, K.,  
 Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gaasterland, T., Gissi, C., King, R., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G.,  
 Quackenbush, J., Schmitt, J., Schmitt, J., Suzuki, P., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Harsh, G., Blake, J., Hoffell, D., Hojuda, N.,  
 Carninci, P., de Bonaldo, M., Brownstein, M. J., Bull, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Holman, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
 Marchionni, J., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,  
 Ringwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Soga, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
 Taya, K. K., Wang, K. H., Weitz, C., Whitaker, C., Wilmink, L.,  
 Wyszewski, A., Yoshida, K., Hasegawa, Y., Kawaji, J., Kotsuki, S.,  
 and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21087660  
 MEDLINE  
 11217851  
 5 (bases 1 to 1364)  
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
 Arai, J., Hagiwara, K., Hono, Y., Brownstein, M., Bull, C.,  
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hagiwara, T.,  
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, P.,  
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
 Kato, H., Kawaji, J., Kojima, Y., Kono, H., Kouda, M., Koyas, S.,  
 Kirihara, K., Matsuyama, T., Miyazaki, A., Nishi, Y., Nomura, K.,  
 Numaoka, K., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
 Saito, H., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
 Schmitt, J., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
 Souabe, Y., Suzuki, H., Tadam, M., Tazawa, A., Takahashi, T.,  
 Tanaka, Y., Tejima, Y., Taya, T., Yamamura, T., Yamanaka, I.,  
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and  
 Hayashizaki, Y.  
 Direct Submission  
 Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@riken.go.jp,  
 URL: http://230-0045-riken.go.jp/). Tel: 81-45-503-9222.









126 ValLeuPheAlaLeuPheSerLysIleuCluArgThrCysCluLeuLeuTyrLeuThrGln 145



```

67 AspSerIleValPheTyAsnSerValPheMetGlnAraLeuLysThrAsnIleLeu 738
|||||
68 GAlIcAlAlAGATATTAACACGGCTCTCATGCAGAGACTCAAAACAATAATTGG 667
|||||
69 GlnTyAlaSerThrArgProPro ThrLeuSerProIleProHisIle-ProArgSe 757
|||||
70 CATTATCTTATTAACCTGGGCTCTCATGAGAGACTGAAANAAATATTTCGATAGT 181
|||||
71 ThrArgProProThrLeuSerProIleProHisIleProArgSerProTyLysPhe 761
|||||
72 TCCACAGAGCTCTTACCTTCGACAAACACACAAATAATGACAAACCTTCAACCT 241
|||||
73 ProSe-SerProLeuArgIleProLysIleLeuSerProIleLysSerPro 781
|||||
74 CTATATTCATATTAAGATTCGAGAGAAATATATATATTAATCAATCAATGAGAG 401
|||||
75 TyLysIleSerGlyGlyLeuProThrProThrLysMetThrProArgSerArgIleLeu 801
|||||
76 TATAAAATTCAGACAGCTGCTGCTCAAAACACACAAATAATGATCCAAAGATCA 461
|||||
77 ValSerIleGlyLysSerIleLeuSerIleLeuSerIleLeuSerIleLeuSer 821
|||||
78 GATCAAAATGGTGCAATTCGGGACTTCGAGAGACTTCGAGAGAAATATAATCA 421
|||||
79 CysAsnSerAspArgValLeuLysAspSerValGlnGlySerAsnProProLysPro 841
|||||
80 TCTAACAGAGAGCTGCTGCTCAAAACACACAAATAATGATCCAAAGATCA 481
|||||
81 LysIleValArgPheAspIleGlySerAspSerValGlnGlySerLysHisLeuPro 861
|||||
82 AAAAAATCACTTTGATATTCAGAGATAGAGAGAAATAGTGGAGATGAAATATCT 541
|||||
83 GlyGlySerLysPheGlnGlyLysLeuAlaGluMetThrSerThrArgThrArgMet 881
|||||
84 GCACATCCCAAAATTCAGCAGAGAAATCCGACAGAAATCACTTCACACACAGAA 601
|||||
85 LysGlyLysMetAspAspSerMetAspThrSerAsnLysGlnGlyLys 897
|||||
86 AGCAGAAATAGATGATAGATGATGATGATGATGATGATGATGATGATGATGAT 649
|||||
RESULT 9
AA763411/C
LOCUS
DEFINITION
Mus musculus
house mouse.
ACCESSION
AA763411.1 GI:2813158
VERSION
AA763411
KEYWORDS
EST
SOURCE
ORGANISM
REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubouque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, H.,
Theising, H., Wylie, T., Lennon, G., Soares, H., Wilson, R. and
Waterston, R.
The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Marra M./Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8507, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty free through LINH; contact the
IMAGE Consortium (info@image.lnh.gov) for further information.
Seq primer: -28m13 rev2 ET from Amerisham
high quality sequence stop: 422.

```

```

67 AspSerIleValPheTyAsnSerValPheMetGlnAraLeuLysThrAsnIleLeu 738
|||||
68 GAlIcAlAlAGATATTAACACGGCTCTCATGCAGAGACTCAAAACAATAATTGG 667
|||||
69 GlnTyAlaSerThrArgProPro ThrLeuSerProIleProHisIle-ProArgSe 757
|||||
70 CATTATCTTATTAACCTGGGCTCTCATGAGAGACTGAAANAAATATTTCGATAGT 181
|||||
71 ThrArgProProThrLeuSerProIleProHisIleProArgSerProTyLysPhe 761
|||||
72 TCCACAGAGCTCTTACCTTCGACAAACACACAAATAATGACAAACCTTCAACCT 241
|||||
73 ProSe-SerProLeuArgIleProLysIleLeuSerProIleLysSerPro 781
|||||
74 CTATATTCATATTAAGATTCGAGAGAAATATATATATTAATCAATCAATGAGAG 401
|||||
75 TyLysIleSerGlyGlyLeuProThrProThrLysMetThrProArgSerArgIleLeu 801
|||||
76 TATAAAATTCAGACAGCTGCTGCTCAAAACACACAAATAATGATCCAAAGATCA 461
|||||
77 ValSerIleGlyLysSerIleLeuSerIleLeuSerIleLeuSerIleLeuSer 821
|||||
78 GATCAAAATGGTGCAATTCGGGACTTCGAGAGACTTCGAGAGAAATATAATCA 421
|||||
79 CysAsnSerAspArgValLeuLysAspSerValGlnGlySerAsnProProLysPro 841
|||||
80 TCTAACAGAGAGCTGCTGCTCAAAACACACAAATAATGATCCAAAGATCA 481
|||||
81 LysIleValArgPheAspIleGlySerAspSerValGlnGlySerLysHisLeuPro 861
|||||
82 AAAAAATCACTTTGATATTCAGAGATAGAGAGAAATAGTGGAGATGAAATATCT 541
|||||
83 GlyGlySerLysPheGlnGlyLysLeuAlaGluMetThrSerThrArgThrArgMet 881
|||||
84 GCACATCCCAAAATTCAGCAGAGAAATCCGACAGAAATCACTTCACACACAGAA 601
|||||
85 LysGlyLysMetAspAspSerMetAspThrSerAsnLysGlnGlyLys 897
|||||
86 AGCAGAAATAGATGATAGATGATGATGATGATGATGATGATGATGATGATGAT 649
|||||
RESULT 9
AA763411/C
LOCUS
DEFINITION
Mus musculus
house mouse.
ACCESSION
AA763411.1 GI:2813158
VERSION
AA763411
KEYWORDS
EST
SOURCE
ORGANISM
REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubouque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, H.,
Theising, H., Wylie, T., Lennon, G., Soares, H., Wilson, R. and
Waterston, R.
The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Marra M./Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8507, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty free through LINH; contact the
IMAGE Consortium (info@image.lnh.gov) for further information.
Seq primer: -28m13 rev2 ET from Amerisham
high quality sequence stop: 422.

```









449 eGlnAsnPheSerLysLeu 445

Qy 344 aAlaArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsuSerAlaSerAsp; 364

QY 344 a lA r q T h r V a l M e t A s n T h r I l e G l n G l n L e u M e t M e t I l e L e u A s u S e r A l a S e r A s p G 364





GenCore version 5.1.4  
Copyright (c) 1993-2003 CompuGen Ltd.

```

c-m protein nucleotide search, using frame_plus.p2n model
Run on: January 17, 2003, 18:27:24 : Search time 2053.03 seconds
(without alignments)
7420.587 Million cell updates/sec

Title: US-09-026-459A-51
Perfect score: 4798
Sequence: 1 MPEKIKRKAIAIAAAAEI.....IMRQKKNISMTSNKEEK 928

Scoring table:
Biosum62: Xgapop 10.0 : Xgapext 0.5
Ygapop 10.0 : Ygapext 0.5
Fgapop 6.0 : Fgapext 7.0
Delop 6.0 : Delext 7.0

Searched: 1615466 seqs, 804774376 residues
Total number of hits satisfying chosen parameters: 32408142

Minimum db seq length: 0
Maximum db seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame.p2n.model -DEV-klh
Q-Z=us-09-026-459a-51
DB=EST_QPMI-fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LXOPEXT=0
UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS-human40.cdi -LIST=45
DBALIGN=200 -THR_SCORE_PCT=THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MCOPI-LOCAL
GAPMI-p2n -WMM-ext -HEADS=128-500 -MINLEN=0 -MAXLEN=200000000
USER US09026459-05N 1 18175 -format 16012003-152138-23322 -NCP0-6 -ICP0-4
No XLPAY No MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
WARN LIM0001 40 THREAS=1 XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
YGAPOP=10 YGAPEXT=0.5 DELOP=6 -DELEXT=7

```

```

Database:
EST:
1: em.est.bm:
2: em.est.bm:
3: em.est.in:
4: em.est.in:
5: em.est.m:
6: em.est.m:
7: em.est.m:
8: em.est.m:
9: qb.est.1:
10: qb.est.2:
11: qb.est.3:
12: qb.est.4:
13: qb.est.4:
14: qb.est.5:
15: em.est.m:
16: em.est.m:
17: qb.qss:
18: em.qss.bm:
19: em.qss.bm:
20: em.qss.pla:
21: em.qss.vrt:
22: em.qss.tun:
23: em.qss.mam:
24: em.qss.mus:
25: em.qss.other:
26: em.qss.pro:
27: em.qss.pro:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1883	97.2	1354	11	AK011246 Mus musculus
2	1551	32.3	1051	13	BM361084
3	1448	40.2	883	14	HQ220275
4	1339	27.9	1001	13	BM475603
5	1254	26.1	727	9	AL597811
6	1246.5	25.8	880	14	BQ220277
7	1208	25.2	845	13	BM454244
8	1111	23.2	1024	13	BM450031
9	1091	22.7	842	9	AI646038
10	1056.5	22.0	914	9	AA764411
11	1043	21.7	792	9	AI564757
12	1037	21.6	695	9	AL596433
13	1030	21.5	694	10	BE282846
14	1016	21.2	624	10	AA584381
15	1013	21.1	871	9	AA764485
16	1004	20.9	922	12	BQ220275
17	978	20.4	588	10	AA368234
18	963.5	20.1	709	9	AI646097
19	962	20.1	736	13	BI1151568
20	959	20.0	594	10	AV715533
21	955	19.9	584	12	HG149050
22	947.5	19.7	734	12	BF144857
23	938	19.5	534	9	AI592790
24	908	18.9	797	10	DE539278
25	903	18.8	615	9	AA595836
26	890	18.5	1154	14	BM362746
27	882	18.4	572	9	AL120941
28	879.5	18.3	744	13	HI520482
29	869.5	18.1	525	12	HG244088
30	858	17.9	557	12	BE941246
31	844	17.6	484	10	BE168095
32	843.5	17.4	646	13	BM490416
33	824	17.2	494	10	AA502887
34	815.5	17.0	522	9	AA072780
35	813	16.9	545	10	HG692240
36	792	16.5	643	13	BI553937
37	784	16.3	2083	13	BM456728
38	751	15.7	453	10	AA503514
39	704	14.7	529	9	AA595709
40	693.5	14.5	472	9	AA595457
41	690	14.4	452	9	AI606529
42	683	14.2	448	12	BF827047
43	681.5	14.2	799	14	CH2404
44	681	14.2	514	12	BF418684
45	667	13.9	651	13	BI696841

## ALIGNMENTS

```

RESULT 1
AK011246
LOCUS
DEFINITION
Mus musculus 10 days embryo whole body cDNA, RIKEN full length
enriched library, clone:260017M22:ref(embryo), full insert
sequence.
ACCESSION
AK011246
VERSION
AK011246.1 GI:12847243
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain:G57H/bJ) 10 days embryo cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2600017M22.
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```

Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 403, 16-44 (1999)  
 99279253  
 MEDLINE  
 10349636  
 PubMed

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 MEDLINE  
 11042159  
 PubMed

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, T., Nishi, K., Kikunishi, Y., Kono, H., Adachi, T., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamashita, M., Matsuda, H., Sakaguchi, S., Kikunishi, Y., Kashiwagi, K.,  
 Fujitake, S., Inoue, K., Terada, Y., Iizawa, M., Ohara, K., Watabiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, T.,  
 Okazaki, Y., Muramatsu, M., Iizawa, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system - 484-format  
 sequencing pipeline with 384 multichannel capillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530413  
 MEDLINE  
 11076861  
 PubMed

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakata, T.,  
 Saito, T., Okazaki, Y., Kikunishi, Y., Bodo, H., Kasukawa, T., Saito, H.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsui, Y., Nishida, T., Paoletti, G.,  
 Quackenbush, J., Schriml, M., Straub, P., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baladracelli, R., Harsh, G., Blake, J., Hoffell, D., Hoshida, N.,  
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Holman, M., Hume, P.A., Kamiya, M., Lee, N.H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,  
 King, P., Ringwald, M., Rodriguez, I., Sarrafian, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Sosa, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,  
 Winkler, B., Boriss, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S.,  
 Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085460  
 MEDLINE  
 11217851  
 PubMed

Aizawa, K., Aizawa, Y., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
 Arai, T., Aizawa, Y., Haldar, P., Hono, H., Brownstein, M., Bull, C.,  
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanada, K.,  
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, P.,  
 Hume, D., Imomari, K., Ishii, Y., Itoh, M., Iizawa, M., Kasukawa, T.,  
 Kato, H., Kawai, T., Kojima, Y., Kono, H., Kouda, M., Koyas, K.,  
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, S.,  
 Numazaki, K., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
 Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tatem, M., Tazawa, A., Takahashi, F.,  
 Tanaka, Y., Tejima, Y., Toyota, Y., Yamamura, T., Yamakata, T.,  
 Yasushita, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
 Hayashizaki, Y.  
 Direct Submission  
 Submitted (10-01-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN) Laboratory for Genome  
 Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-1-1 Hirosaki, Saitama, 350-0505, Japan.  
 Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
 URL: <http://genome-res.riken.go.jp/>, Tel: 81-45-503-9226.















1000000

1000000

1000000





tel: 65-11-2704922  
 Fax: 65-11-2707001  
 Email: asinspcond@indiana.edu  
 This sequence was derived from the FAPSP/LLRP Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.indiana.edu/~llrp/scripts/gene.html) EST ID: 100642-210  
 pub 011-09-01-2000 02 21541-1

Seq primer: puc 18 forward  
 High quality sequence start: 58  
 High quality sequence stop: 669.  
 Location/Qualifiers

## FEATURES

1..694  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BT0642"  
 /clone\_stage="Adult"  
 /note="Organ: Breast; Vector: puc18; Site: 1; SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ONESTES PCR (0.8 kb) from Patient application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 211 a 129 c 127 g 226 t 1 others

## BASE COUNT

Alignment Scores:  
 Pred. No.: 1,42e-92 Length: 694  
 Score: 1030.00 Matches: 297  
 Percent Similarity: 96.33% Conservative: 3  
 Best Local Similarity: 94.45% Mismatches: 2  
 Query Match: 21.47% Indels: 7  
 Gaps: 0

US-09-026-459a-51 (1-928) x BE082846 (1-694)

405 AsnCysHisValAsnProGysGlnSerIleLeuGysArgValLysAspIleClyTyrIle 425  
 406 CACTGACACATAT-CCAAAGCAAGATATATCTGAGAGAGTGAAGATATAGATATATC 618  
 407 PheLysGlnLysPheAlaValGlyGlnGlyCysValGluIleLeuSerGlnArg 445  
 408 TTTTAAACACAAAATTGCTACAGCTGAGGACACAGGCTGAGGAAATTCGAAACACAGCA 558  
 409 TyrLysLeuGlyValArgLeuTyrTyrArgValMetGlnSerMetLeuLysSerGluGln 465  
 410 TACAAATTTGAT-ATTCGCTTATTACGAGTAATGGAATATGCTTTAAATATCAAGCA 498  
 411 GlnArgLeuSerIleGlnAsnProGysGlnSerIleLeuGysArgValLysAspIleClyTyrIle 485  
 412 GAAGCATATAT-ATTTAAATTTTATG-AAAATTCGACATGAAATTTTTCATATATCT 438  
 413 LeuLeuAlaCysAlaLeuGlnValValMetAlaThrTyrSerArgSerThrSerGlnAsn 505  
 414 TTAAGCGGCGGCTCTCTCAAGGCTGAAAGGCAATATATGCAAGTATATATATATAT 378  
 415 LeuAsnSerGlyThrAspLeuSerPheProTyrIleLeuAsnValLeuAsnLeuLysAla 525  
 416 GTTATCTGGACACATAGTCTTCTGATGATCTGATGATGATGATGATGATGATGATG 318  
 417 PheAspHisThrLysValIleGlnSerPheIleLysAlaGluGlyAsnLeuThrArgGlu 545  
 418 TTTTATTTTAT-AAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 258  
 419 MetIleLysHisLeuGlnArgCysIleHisArgTlleMeIleGlnSerLeuAlaTrpLeuSer 565  
 420 ALGAAACAACTTACACACGACGACAAATGCAATGCAATGCAATGCAATGCAATGCA 198  
 421 AspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGluGlyProThrAspHis 585  
 422 GATTCACCTTATTATTTTATTATTAAT-AAAATGATGATGATGATGATGATGATGAT 138  
 423 LeuGlnSerAlaLysProLeuAsnLeuProGlnGluAsnLeuAsnHisThrAlaLeuAspMe 605

137 CTGGAATCTGCTTCCCTTAACTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT 78  
 605 tTyrLeuSerIleValArgSerProLysLysLysLysSerThrThrArgVal 622  
 77 CTATCT 26

## RESULT 14

AW583181 623 bp mRNA linear EST 13-MAR-2002  
 Locus: t11h09.y1 Human Pancreatic Islets Homo sapiens cDNA clone  
 IMAGE:5638049 5' similar to gb:M15430 RELINCHI/AS/OMA-AS/OMA/1A/BD  
 PROTEIN (HUMAN): gb:M26361 Mouse retinoblastoma susceptibility  
 protein (MUSP); mRNA sequence.

## ACCESSION

AW583181 1 31:250125

## VERSION

EST

## KEYWORDS

human

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
 1 (bases 1 to 623)  
 Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,  
 Weller,T., Martin,J., Ristain,A., Schmitt,A., Theising,H., Ritter  
 R., Piro,L., Bennett,J., Gajewski,M., Gibbons,M., McCann,P., Cole  
 R., Tsagaris,W., Williams,J., Jackson,Y. and Howers,Y.  
 WashU-Harvard Pancreas EST Project

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@bioph.harvard.edu

Libraries were constructed by Dr. Douglas Melton  
 DNA sequencing by Washington University Genome Sequencing Center  
 For information on obtaining a clone please contact: Juliana Brown  
 (brownstas.harvard.edu)  
 This sequence now available from the IMAGE consortium, for clone  
 orders contact: info@image.llnl.gov  
 Seq primer: 40RP from Gibco  
 High quality sequence stop: 410.  
 Location/Qualifiers

## FEATURES

## Source

1..623  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:5638049"  
 /clone\_stage="Adult"  
 /issue\_type="Islets of Langerhans"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B"  
 /note="Organ: Pancreas; Vector: pSM0811; Site: 1; Not 1;  
 Site 2: Sal 1; Library constructed using SuperScript  
 Plasmid Library kit (Life technologies) cDNA made by  
 cloning of primary size selected by column fractionation:  
 average insert size 1.08 kb. Primary library,  
 unamplified."  
 234 a 93 c 109 g 186 t 1 others

## BASE COUNT

## ORIGIN

## Alignment Scores:

Pred. No.: 3.05e-91 Length: 624  
 Score: 1016.00 Matches: 201  
 Percent Similarity: 98.07% Conservative: 2  
 Best Local Similarity: 97.10% Mismatches: 3  
 Query Match: 21.18% Indels: 1  
 Gaps: 0

US-09-026-459a-51 (1-928) x AW583181 (1-623)

## QY

271 ThrArgIleCysValLeuGysLysGlnHisGlnCysAsnIleCysGlyValLysAsn 290





```

QY 555 isAqilleMetGluSerLeuAlaTrpLeuSerAspSerLeuLeuPheAspGluIleLysG 575
1b 264 ACGAAACATGGAAACCCCTGGCAAGGCTTTCAGATTCACCTTTATTTGATCTCATTAAGC 205
QY 575 InSerLysAspArgGluGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuP 595
1b 264 ASTCCAAAGATGCAAGAGACCT---GATAAAGTTTAAAGTCTTTCTCTCTAAGATCTGAAAGA 148
QY 595 rGleuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValAlaArgSerProLysL 615
1b 147 CTCTCAGAGGTAAAGCATATAGTGGAGAGAGATAGTATCTTTCTCTCTAAGATCTGAAAGA 88
QY 615 ysLysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerA 635
1b 67 AAATAACATCCATACACCGTGTAATTCCTGCAAAATACAGAGACACAAGCAGCCTCAG 28
QY 635 laPheGlnThrGlnLysProLeuLys 643
1b 27 GCTTCATATAGTCAGAGGATTGAAA 2

```

Search completed: January 19, 2003, 05:07:40  
 Job time: 2687.11 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Computer Ltd.

EM nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:20:22, Search time 316.42 seconds  
(without alignments)  
18469,506 Million cell updates/sec

Hit(s): 08-09-026-459A-50

Partial Score: 4554  
Sequence: 1 GGGTCATGGGCGCAAAAC.....AAATGAGGATTATGATAGT 4554

Scoring table: IDENTITY NC  
Gap: 10.0, Export: 1.0

Searched: 16154666 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

Listed first 45 summaries

Database: EST:

- 1: em\_estb1.\*
- 2: em\_estb1.\*
- 3: em\_estb1.\*
- 4: em\_estb1.\*
- 5: em\_estb1.\*
- 6: em\_estb1.\*
- 7: em\_estb1.\*
- 8: em\_estb1.\*
- 9: em\_estb1.\*
- 10: em\_estb1.\*
- 11: em\_estb1.\*
- 12: em\_estb1.\*
- 13: em\_estb1.\*
- 14: em\_estb1.\*
- 15: em\_estb1.\*
- 16: em\_estb1.\*
- 17: em\_estb1.\*
- 18: em\_estb1.\*
- 19: em\_estb1.\*
- 20: em\_estb1.\*
- 21: em\_estb1.\*
- 22: em\_estb1.\*
- 23: em\_estb1.\*
- 24: em\_estb1.\*
- 25: em\_estb1.\*
- 26: em\_estb1.\*
- 27: em\_estb1.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Length	ID	Description
1	908.4	27.2	104	11	AK011246 Mus musc
2	910.2	25.6	1023	13	BM450031 AGENCGURT
3	909.4	25.6	1051	13	BM450031 AGENCGURT
4	856	24.1	1081	14	BM450031 AGENCGURT
5	844.8	23.9	981	14	BM450031 AGENCGURT
6	786	22.1	1001	13	BM450031 AGENCGURT

7	752.4	21.2	1122	13	BM450031 AGENCGURT
8	725.4	20.4	727	9	AL5597811 DKZP4146
9	720.6	20.3	886	14	BM450031 AGENCGURT
10	699.6	19.7	823	12	BM450031 AGENCGURT
11	698.2	19.5	792	9	AL5597811 DKZP4146
12	692.2	19.5	767	12	BM450031 AGENCGURT
13	691.8	19.5	972	13	BM450031 AGENCGURT
14	678	19.1	945	13	BM450031 AGENCGURT
15	660	18.6	672	13	BM450031 AGENCGURT
16	644.6	18.1	797	10	BM450031 AGENCGURT
17	642.2	18.1	681	12	BM450031 AGENCGURT
18	617	17.4	618	13	BM450031 AGENCGURT
19	614	17.3	694	10	BM450031 AGENCGURT
20	605.2	17.0	623	10	BM450031 AGENCGURT
21	605	17.0	605	9	AL5597811 DKZP4146
22	595.4	16.8	922	12	BM450031 AGENCGURT
23	581.6	16.4	832	9	AL5597811 DKZP4146
24	580	16.3	580	9	AL5597811 DKZP4146
25	575.2	16.2	914	9	AA764411 VW54404.1
26	566.4	15.9	588	10	AA764411 VW54404.1
27	558	15.7	871	9	AA764411 VW54404.1
28	530.2	14.9	598	10	AA764411 VW54404.1
29	530	14.9	539	9	AA764411 VW54404.1
30	528	14.9	572	9	AA764411 VW54404.1
31	516.8	14.5	597	12	BM450031 AGENCGURT
32	513.8	14.5	1154	14	BM450031 AGENCGURT
33	513.2	14.4	594	10	AA764411 VW54404.1
34	512.2	14.4	734	12	BM450031 AGENCGURT
35	509.2	14.3	709	9	AA764411 VW54404.1
36	505.2	14.2	736	13	BM450031 AGENCGURT
37	493.8	13.9	525	9	AA764411 VW54404.1
38	486.4	13.7	584	12	BM450031 AGENCGURT
39	484.4	13.6	493	10	AA764411 VW54404.1
40	481.4	13.5	484	10	AA764411 VW54404.1
41	480.2	13.5	659	13	BM450031 AGENCGURT
42	476.6	13.4	2084	13	BM450031 AGENCGURT
43	475.4	13.4	615	9	AA764411 VW54404.1
44	474.4	13.3	734	13	BM450031 AGENCGURT
45	459.2	12.9	497	12	BM450031 AGENCGURT

ALIGNMENTS

RESULT 1	AK011246	1464 bp	mkna	Insert	Hit 19 JAN 2002
LOCUS	Mus musculus 10 days embryo whole body cDNA, RIKEN full length				
DEFINITION	enriched library, clone:2690317M22, rat fibroblastoma 1, full insert				
ACCESSION	AK011246				
VERSION	AK011246.1	GI:12847243			
KEYWORDS	HTC; CAP trapper				
SOURCE	Mus musculus (strain:57BL/6J) 10 days embryo cDNA to mRNA, clone:lib-RIKEN full-length enriched mouse cDNA library, clone:2690317M22				
ORGANISM	Mus musculus				
REFERENCE	1 Fukuyama, M.; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PubMed	10349636				
REFERENCE	2				
AUTHORS	Carninci, P.; Shibata, Y.; Hayashizaki, Y.; Sasaki, K.; Itoh, M.; Kono, H.; Kuzaki, Y.; Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PubMed	11042159				





















sequences: 5' CAGCGGCAATTAAGGTC 4' and 4' adaptor sequence:  
5' ATTCTACAGGTCAGAGTCGTCGATAG-dl (30)NN 3' (where N = A,  
C, G, or T). Average insert size 1.9  
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH MGC library.\*

```

BASE COUNT      233 a   125 c   140 g   279 t
CGTGTG

Query Match      19.5%; Score 692.2; DB 12; Length 767;
Best Local Similarity 97.1%; Pred. No. 2,66 125;
Matches 74; Conservative 0; Mismatches 18; Indels 4; Gaps 4;

QY 2762 GATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 2821
DB 1 GATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 60

QY 2822 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 2881
DB 61 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 120

QY 2882 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 2941
DB 121 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 180

QY 2942 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3001
DB 181 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 240

QY 3002 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3061
DB 241 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 400

QY 3062 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3121
DB 401 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 360

QY 3122 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3181
DB 461 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 420

QY 3182 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3241
DB 421 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 480

QY 3242 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3300
DB 481 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 540

QY 3300 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3360
DB 541 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 600

QY 3360 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3420
DB 601 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 659

QY 3421 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3480
DB 660 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 718

QY 3481 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 3540
DB 719 TATGATATAGTCAAGCAAGCAAGCAAGTATGAGATCTTGCTGACAACTGTG 767

PEP011114      972 bp      mRNA      linear      EST 05-FEB-2002
LOCUS          RM466194
DEFINITION     AGEN0907L_6456/73 NIH MGC_92 Homo sapiens cDNA clone IMAGE:5576939
ACCESSION     RM466194

```

```

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

```

```

RM466194.1  GI:18515236
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 972)
NIH MGC http://mimc.nhlbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgs@nsl.nhlbi.nih.gov
Tissue Procurement: AICC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2429 row: j column: 12
High quality sequence stop: 621.
Location/Qualifiers
1..972
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH MGC_92"
/issue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="organ: testis; Vector: pCMV Sport6; Site 1: Not1;
Site 2: SalI; Cloned unidirectionally; origin of primed.
Average insert size 2.5 kb. Library enriched for
full length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
2 others
BASE COUNT      294 a   184 c   169 g   423 t
ORIGIN

```

```

Query Match      19.5%; Score 691.8; DB 13; Length 972;
Best Local Similarity 96.7%; Pred. No. 3,16 125;
Matches 728; Conservative 0; Mismatches 19; Indels 6; Gaps 2;

QY 2698 CAGCAGAAATGGGCAAAAGACTTCTACGCAACACAGCAAGCAAGCAAGCAAG 2757
DB 1 CAGCAGAAATGGGCAAAAGACTTCTACGCAACACAGCAAGCAAGCAAGCAAGCAAG 60

QY 2758 GATAGCAATGATATCTCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2817
DB 61 GATAGCAATGATATCTCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 120

QY 2818 TCTGTACACCTCTGCAATTCATCTGCTCTGCAACATGATGATGATGATGAT 2877
DB 121 TCTGTACACCTCTGCAATTCATCTGCTCTGCAACATGATGATGATGATGATGAT 180

QY 2878 TGTTATGTCGACATTAATATCTCTGCAACATGATGATGATGATGATGATGAT 2937
DB 181 TGTTATGTCGACATTAATATCTCTGCAACATGATGATGATGATGATGATGAT 240

QY 2938 ATTGTTGGGTGATCTTCTAAGGCACTTGAATGTTTATGTTTATTTATTTAT 2997
DB 241 ATTGTTGGGTGATCTTCTAAGGCACTTGAATGTTTATGTTTATTTATTTATTT 600

QY 2998 AAAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3057
DB 401 AAAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460

QY 3058 TAAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3117
DB 361 TAAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 3118 AGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3177
DB 421 AGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

```



```

/Ab xref "taxon:9606"
/clone IMAGE:5402622
/clone lib "NIH_M07_90"
/issue type "adenocarcinoma, cell line"
/lab host "pH10H (phage-resistant)"
/notes "organ: liver; Vector: pCMV-SPORT6; Site_1: Nott;
Site_2: Salt; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full length clones and constructed by life technologies.
Note: this is a NIH MCC library."
BASE COUNT      228 a   124 c   144 g   186 t
          49101N

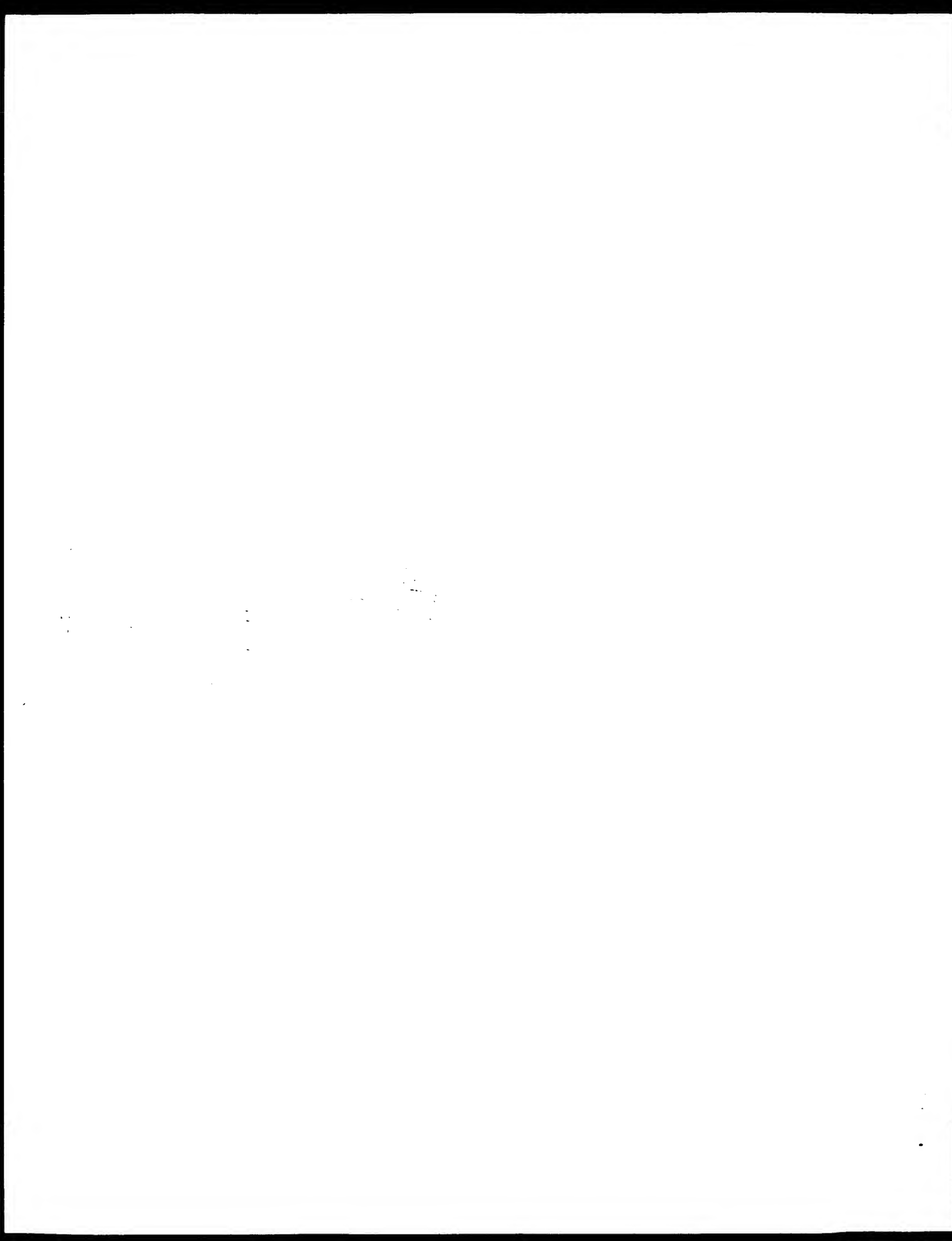
```

```

Query Match      18.6%; Score 660; DB 13; Length 672;
Best local Similarity 99.9%; Prod. No. 5,3e 119;
Matches 671; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2420 ATATTTCAGCTGGAGATGATATATAAATTTTCAGAAAGTGTGCAACAGCAACAAAAA 2479
LB 1 AATTTTCAGCTGGAGATGATATATAAATTTTCAGAAAGTGTGCAACAGCAACAAAAA 60
QY 2480 GCAATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 2539
LB 1 GCAATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 120
QY 2540 TCCAGAAATATATCATATGATATGTAACAGGAGCTGTGTCAAAAAGTGTGAAG 2599
LB 1 TCCAGAAATATATCATATGATATGTAACAGGAGCTGTGTCAAAAAGTGTGAAG 180
QY 2600 GAAATGAGCTGTGTCAAAAAGTGTGTGTGATATTCAGAGATCATATGAAG 2659
LB 1 GAAATGAGCTGTGTCAAAAAGTGTGTGTGATATTCAGAGATCATATGAAG 240
QY 2660 CATATGCAATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 2719
LB 1 CATATGCAATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 300
QY 2720 CTTCATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 2779
LB 1 CTTCATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 360
QY 2780 AGCAATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 2839
LB 1 AGCAATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 420
QY 2840 GCTCTGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 2899
LB 1 GCTCTGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 480
QY 2900 CTTCATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 2959
LB 1 CTTCATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 540
QY 2960 CATATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 3018
LB 1 CATATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 600
QY 3019 CATATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 3078
LB 1 CATATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 660
QY 3079 ATATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 3138
LB 1 ATATGCAAGATCAAAATCTAGCAATCAATGGGCAATCATGGGAGCTGGAGAGT 672

```

Search completed: January 18, 2003, 04:11:49  
 Feb time : 3148.5 secs





Genfore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:29:22 / Search time 2663.48 seconds  
(without alignments)  
1846% 506 Million cell updates/sec

Filter: US 09 026-459A 34

Perfect score: 4266  
Sequences: 1 GCAATCAATGCGGGGAA.....AAATGAGTAATTAATCACT 3266

Scoring table:

IDENTITY: 100

Gapop 10.0 / Gapext 1.0

Searches: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 42408142

Minimum db seq length: 0

Maximum db seq length: 20000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database: EST

1: cm\_estbac\*  
2: cm\_estbac\*  
3: cm\_estbac\*  
4: cm\_estbac\*  
5: cm\_estbac\*  
6: cm\_estbac\*  
7: cm\_estbac\*  
8: cm\_estbac\*  
9: cm\_estbac\*  
10: cm\_estbac\*  
11: cm\_estbac\*  
12: cm\_estbac\*  
13: cm\_estbac\*  
14: cm\_estbac\*  
15: cm\_estbac\*  
16: cm\_estbac\*  
17: cm\_estbac\*  
18: cm\_estbac\*  
19: cm\_estbac\*  
20: cm\_estbac\*  
21: cm\_estbac\*  
22: cm\_estbac\*  
23: cm\_estbac\*  
24: cm\_estbac\*  
25: cm\_estbac\*  
26: cm\_estbac\*  
27: cm\_estbac\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	919.2	27.9	1623	BM450031	BM450031 AGENCOURT
2	909.4	27.8	1051	BM561083	BM561083 AGENCOURT
3	866	26.2	1981	BQ43221	BQ43221 AGENCOURT
4	844.8	25.9	881	BQ220275	BQ220275 AGENCOURT
5	786	24.1	1001	BM475603	BM475603 AGENCOURT
6	777.8	23.8	1363	AK011246	AK011246 Mus muscu

RESULT 1  
BM450031  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT  
FEATURES

1024 bp mRNA linear EST 09-SEP-2002  
AGENCOURT 4392544 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE528037  
5' mRNA sequence  
BM450031 GI:18499071  
BM450031  
EST  
human  
Homo sapiens  
Enkaryota; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 1024)  
NIH-MGC http://mgi.nih.gov/2/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: rga@psh-rmail.nih.gov  
Tissue Procurement: ATCC/CRL/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at:  
http://image.liln.gov  
Flate, LILN2234 row 3 column 22  
High quality sequence stop: 643  
location/Qualifiers  
1. 1023

#### ALIGNMENTS

1122 13 BM545881  
727 9 AL597811  
820 14 BQ222227  
823 12 BG610661  
792 9 AL563957  
767 12 BQ570156  
972 13 BM466194  
835 13 BM453724  
672 13 B1B68459  
797 12 BE592738  
621 12 BG616219  
618 13 BM264009  
694 10 BE082846  
623 10 AW583141  
605 9 AL599633  
922 12 BQ353543  
580 9 AL599105  
914 9 AA763411  
588 10 AA763485  
871 9 AA763485  
598 10 AA763485  
539 9 AL692790  
572 9 AL120941  
597 12 BE931236  
594 10 AV715533  
734 12 BF344857  
736 13 B1151568  
525 9 AA258255  
584 12 BG149950  
493 10 AW502887  
484 10 BE168095  
659 13 B1B21116  
2083 13 BM456728  
734 13 B1528682  
497 12 BE944784  
457 9 AL094215  
732 9 AL598766  
869 12 BE865749  
447 9 AT082179



```

101 121 AATTATAGCAAAATTCTTAATGAAATTTTTCATATATGCTTTTATTGAGGTGGGCTCTT 189
102 1192 GAGTGTGTAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1251
103 181 GAGTGTGTAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 240
104 1252 TGGCTTTTGGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1411
105 241 TGGCTTTTGGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 300
106 1312 AATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1471
107 301 AATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 360
108 1372 GAGTGTGTAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1431
109 401 AATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 420
110 1442 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1491
111 421 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 480
112 1492 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1551
113 481 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 540
114 1552 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1611
115 541 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 600
116 1612 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1671
117 601 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 660
118 1672 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1731
119 661 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 720
120 1732 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1791
121 721 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 780
122 1792 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1851
123 781 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 840
124 1852 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1910
125 841 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 900
126 1911 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1962
127 901 GTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 953
128
129 RESULT 4
130 HQ4.9221
131 LOCUS
132 DEFINITION
133 VERSION
134 KEYWORDS
135 SOURCE
136 ORGANISM
137
138 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
139 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
140 1 (bases 1 to 1081)
141 NIH_MGC
142 http://www.ncbi.nlm.nih.gov/
143
144 Unpublished (1999)
145 Contact: Robert Strausberg, Ph.D.
146 Email: rstraus@mail.nih.gov

```

Tissue Procurement: Atrac  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (HLNI)  
DNA Sequencing by: Amersham Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/HLNI at:  
<http://image.llnl.gov>  
Plate: LLAM13211 Row: k Column: 01  
High quality sequence step: 742  
Location/Qualifiers  
1. 1081  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6015480"  
/clone\_lib="NIH\_MGC\_92"  
/tissue\_type="Embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage resistant)"  
/note="Organ: testis; Vector: pCMV-Sp-RP6; Site: 1; Not 1;  
Site\_2: Salt; Cloned unidirectionally; cloned at primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by life technologies.  
Note: this is a NIH\_MGC library."

BASE COUNT 337 a 190 c 183 g 371 t

# ORIGIN

Query Match 26.2% Score 896; DB 14; Length 1081;  
Best Local Similarity 99.1% Pred. No. 1,26-154;  
Matches 892; Conservative 0; Mismatches 5; Indels 3; Gaps 4;

```

QY 2368 GAAGCAGATGAGTAAATATCTCCAGATAGTCTAAATTTAGAGCAAAATGGCAAA 2427
DB 1 GAAGCAGATGGAAGTAAATATCTCCAGATAGTCTAAATTTAGAGCAAAATGGCAAA 60
QY 2428 ATCAGCTCTACCTCAACATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2487
DB 61 ATGACTTCTACTCAACATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 120
QY 2488 AACAGAGGAGAAATGAGATCTCAGATCTTGTGGACACTGTGTACACTGTGATT 2547
DB 121 AACAGAGGAGAAATGAGATCTCAGATCTTGTGGACACTGTGTACACTGTGATT 180
QY 2548 CATCTCTCTACAGATGAGTGTATATTTTCCAGATCTTGTGTATATATATTA 2607
DB 181 CATCTCTCTACAGATGAGTGTATATTTTCCAGATCTTGTGTATATATATTA 240
QY 2608 ATATCTCTACAGATCTTGTGGATATATAATGAGATATATAATGAGATATATTA 2667
DB 241 ATATCTCTACAGATCTTGTGGATATATAATGAGATATATAATGAGATATATTA 400
QY 2668 AAGCAGATGAAATGAGATCTTGTGTATATTTATATATATATATATATATATATAT 2727
DB 301 AAGCAGATGAAATGAGATCTTGTGTATATTTATATATATATATATATATATATAT 460
QY 2728 TGCCATTTAAAAAGTTGAGCAGATTTGCTCTTTTCCAGATATATAATGAGATATTA 2787
DB 361 TGCCATTTAAAAAGTTGAGCAGATTTGCTCTTTTCCAGATATATAATGAGATATTA 420
QY 2788 TGCCATTTAAAAAGTTGAGCAGATTTGCTCTTTTCCAGATATATAATGAGATATTA 2847
DB 421 TGCCATTTAAAAAGTTGAGCAGATTTGCTCTTTTCCAGATATATAATGAGATATTA 480
QY 2848 TGCCATTTAAAAAGTTGAGCAGATTTGCTCTTTTCCAGATATATAATGAGATATTA 2907
DB 481 TGCCATTTAAAAAGTTGAGCAGATTTGCTCTTTTCCAGATATATAATGAGATATTA 540
QY 2908 TTTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2967
DB 541 TTTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 600
QY 2968 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3027
DB 601 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

```









```

FEATURES
Source Berlin:Charlottenburg GPPMANY: Email:elone@rcp.f-dp
Location/Qualifiers
1..727
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFp31400B16"
/clone_lib "313 (synonym: hlec2)"
/dev_stage "adult"
/lab_host="bml04"
/Note="Vector: pTriPlex2, Site_1: SfilA; Site_2: SfilB;
cDNA-collection"
BASE COUNT 249 a 161 c 130 t 198 t
ORIGIN
Query Match 22.1%; Score 725.4; DB 9; Length 727;
Best Local Similarity 99.9%; Pred. No. 1.6e-129;
Matches 726; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 144 GAAATCATAGCTCTTCAATGTTTTATATAAAAAGTATAGGCTATGCTGAGCT 1703
DB 1 GAAATCATAGCTCTTCAATGTTTTATATAAAAAGTATAGGCTATGCTGAGCT 60
QY 154 AAATACATTTTGACAGCGCTGCTAGCTGAGCAGCAGAGATAGACATATCAGCAG 1763
DB 61 AAATACATTTTGAGAGCGCTGCTAGCTGAGCAGCAGAGATAGACATATCAGCAG 120
QY 164 GCTTTTCAGCACAGCTGCGAGAGATGATGAATCATCAGACAGCAGCAATTCGACCA 1823
DB 121 GCTTTTCAGCACAGCTGCGAGAGATGATGAATCATCAGACAGCAGCAATTCGACCA 180
QY 184 AATTATATGTTGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1883
DB 181 AATTATATGTTGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 184 AATCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1943
DB 241 AATCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 144 TTGATCAAAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2003
DB 431 TTGATCAAAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 360
QY 2004 GAGATCAAAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2063
DB 361 GAGATCAAAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 420
QY 2064 AGCTGATTTGCTGGAAGCGCTTAAAGTTCTGATGATGATGATGATGATGATGATGAT 2123
DB 421 AGCTGATTTGCTGGAAGCGCTTAAAGTTCTGATGATGATGATGATGATGATGATGAT 480
QY 2124 GAACATGATATATTCACGCGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2183
DB 481 GAACATGATATATTCACGCGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540
QY 2184 AA'AAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2243
DB 641 AA'AAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 2244 TGAAGATTTGCTGGAAGCGCTTAAAGTTCTGATGATGATGATGATGATGATGATGAT 2303
DB 601 TGAAGATTTGCTGGAAGCGCTTAAAGTTCTGATGATGATGATGATGATGATGATGAT 660
QY 2304 TGAAGATTTGCTGGAAGCGCTTAAAGTTCTGATGATGATGATGATGATGATGATGAT 2363
DB 661 TGAAGATTTGCTGGAAGCGCTTAAAGTTCTGATGATGATGATGATGATGATGATGAT 720
QY 2364 AGATGAA 2370
DB 721 AGATGAA 727
RESULT 9
BQ222227

```





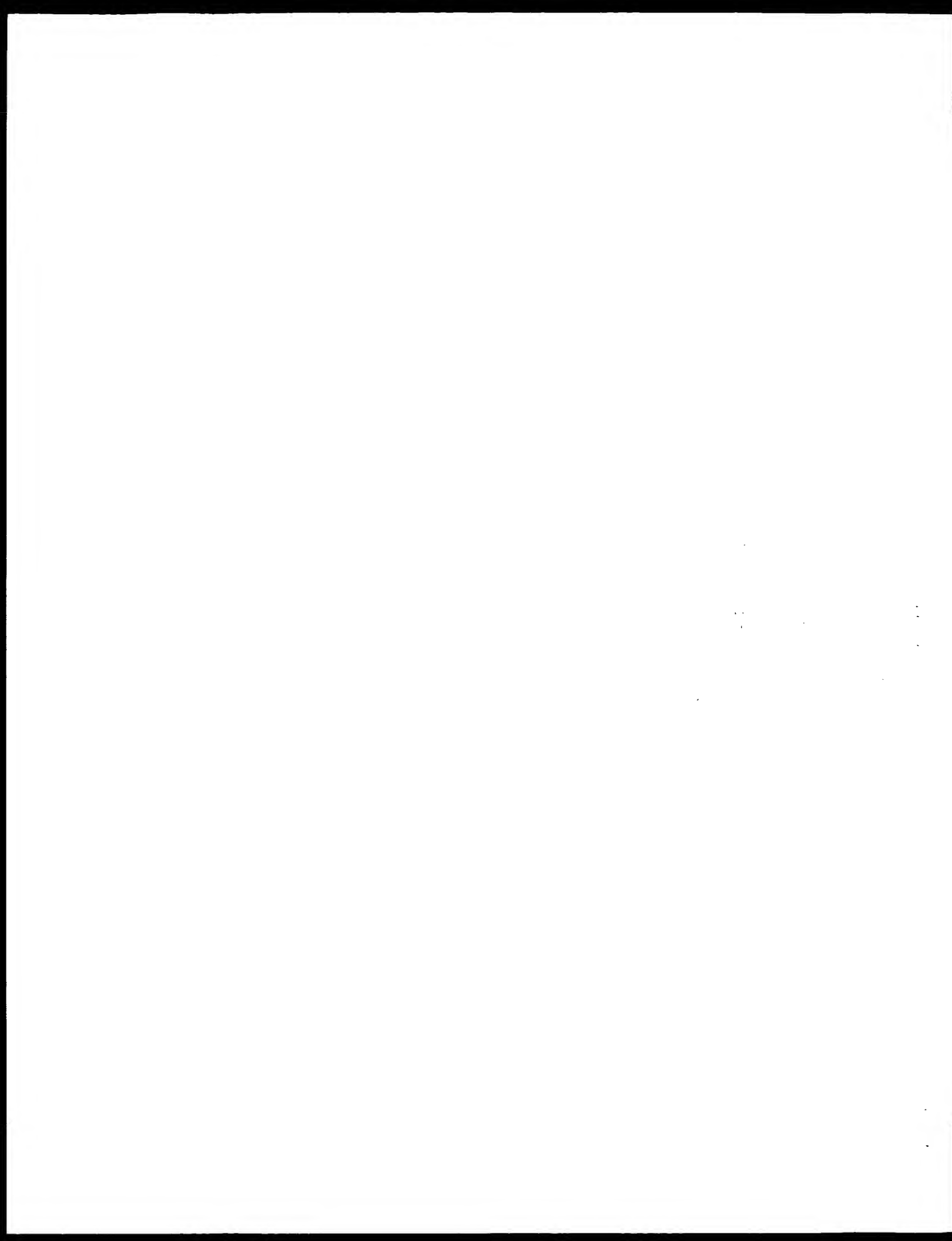






27 2192 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2251  
DB 61 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 120  
27 2252 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2311  
DB 121 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 180  
27 2312 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2371  
DB 161 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 240  
27 2372 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2431  
DB 241 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 300  
27 2432 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2491  
DB 301 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 360  
27 2492 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2551  
DB 361 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 420  
27 2552 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2611  
DB 421 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 480  
27 2612 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2671  
DB 481 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 540  
27 2672 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2730  
DB 541 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 600  
27 2731 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 2790  
DB 601 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT 660  
27 2791 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT  
DB 661 TCAATCAAAATCAAGAAATTTAGTATCAATTTGTTGAATCATTTGGGAAATTTGTGAGAAAT

Search completed: January 14, 2004, 04:07:55  
Job time: 2894.96 secs











```

1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622
2623
2624
2625
2626
2627
2628
2629
2630
2631
2632
2633
2634
2635
2636
2637
2638
2639
2640
2641
2642
2643
2644
2645
2646
2647
2648
2649
2650
2651
2652
2653
2654
2655
2656
2657
2658
2659
2660
2661
2662
2663
2664
2665
2666
2667
2668
2669
2670
2671
2672
2673
2674
2675
2676
2677
2678
2679
2680
2681
2682
2683
2684
2685
2686
2687
2688
2689
2690
2691
2692
2693
2694
2695
2696
2697
2698
2699
2700
2701
2702
2703
2704
2705
2706
2707
2708
2709
2710
2711
2712
2713
2714
2715
2716
2717
2718
2719
2720
2721
2722
2723
2724
2725
2726
2727
2728
2729
2730
2731
2732
2733
2734
2735
2736
2737
2738
2739
2740
2741
2742
2743
2744
2745
2746
2747
2748
2749
2750
2751
2752
2753
2754
2755
2756
2757
2758
2759
2760
2761
2762
2763
2764
2765
2766
2767
2768
2769
2770
2771
2772
2773
2774
2775
2776
2777
2778
2779
2780
2781
2782
2783
2784
2785
2786
2787
2788
2789
2790
2791
2792
2793
2794
2795
2796
2797
2798
2799
2800
2801
2802
2803
2804
2805
2806
2807
2808
2809
2810
2811
2812
2813
2814
2815
2816
2817
2818
2819
2820
2821
2822
2823
2824
2825
2826
2827
2828
2829
2830
2831
2832
2833
2834
2835
2836
2837
2838
2839
2840
2841
2842
2843
2844
2845
2846
2847
2848
2849
2850
2851
2852
2853
2854
2855
2856
2857
2858
2859
2860
2861
2862
2863
2864
2865
2866
2867
2868
2869
2870
2871
2872
2873
2874
2875
2876
2877
2878
2879
2880
2881
2882
2883
2884
2885
2886
2887
2888
2889
2890
2891
2892
2893
2894
2895
2896
2897
2898
2899
2900
2901
2902
2903
2904
2905
2906
2907
2908
2909
2910
2911
2912
2913
2914
2915
2916
2917
2918
2919
2920
2921
2922
2923
2924
2925
2926
2927
2928
2929
2930
2931
2932
2933
2934
2935
2936
2937
2938
2939
2940
2941
2942
2943
2944
2945
2946
2947
2948
2949
2950
2951
2952
2953
2954
2955
2956
2957
2958
2959
2960
2961
2962
2963
2964
2965
2966
2967
2968
2969
2970
2971
2972
2973
2974
2975
2976
2977
2978
2979
2980
2981
2982
2983
2984
2985
2986
2987
2988
2989
2990
2991
2992
2993
2994
2995
2996
2997
2998
2999
3000
3001
3002
3003
3004
3005
3006
3007
3008
3009
3010
3011
3012
3013
3014
3015
3016
3017
3018
3019
3020
3021
3022
3023
3024
3025
3026
3027
3028
3029
3030
3031
3032
3033
3034
3035
3036
3037
3038
3039
3040
3041
3042
3043
3044
3045
3046
3047
3048
3049
3050
3051
3052
3053
3054
3055
3056
3057
3058
3059
3060
3061
3062
3063
3064
3065
3066
3067
3068
3069
3070
3071
3072
3073
3074
3075
3076
3077
3078
3079
3080
3081
3082
3083
3084
3085
3086
3087
3088
3089
3090
3091
3092
3093
3094
3095
3096
3097
3098
3099
3100
3101
3102
3103
3104
3105
3106
3107
3108
3109
3110
3111
3112
3113
3114
3115
3116
3117
3118
3119
3120
3121
3122
3123
3124
3125
3126
3127
3128
3129
3130
3131
3132
3133
3134
3135
3136
3137
3138
3139
3140
3141
3142
3143
3144
3145
3146
3147
3148
3149
3150
3151
3152
3153
3154
3155
3156
3157
3158
3159
3160
3161
3162
3163
3164
3165
3166
3167
3168
3169
3170
3171
3172
3173
3174
3175
3176
3177
3178
3179
3180
3181
3182
3183
3184
3185
3186
3187
3188
3189
3190
3191
3192
3193
3194
3195
3196
3197
3198
3199
3200
3201
3202
3203
3204
3205
3206
3207
3208
3209
3210
3211
3212
3213
3214
3215
3216
3217
3218
3219
3220
3221
3222
3223
3224
3225
3226
3227
3228
3229
3230
3231
3232
3233
3234
3235
3236
3237
3238
3239
3240
3241
3242
3243
3244
3245
3246
3247
3248
3249
3250
3251
3252
3253
3254
3255
3256
3257
3258
3259
3260
3261
3262
3263
3264
3265
3266
3267
3268
3269
3270
3271
3272
3273
3274
3275
3276
3277
3278
3279
3280
3281
3282
3283
3284
3285
3286
3287
3288
3289
3290
3291
3292
3293
3294
3295
3296
3297
3298
3299
3300
3301
3302
3303
3304
3305
3306
3307
3308
3309
3310
3311
3312
3313
3314
3315
3316
3317
3318
3319
3320
3321
3322
3323
3324
3325
3326
3327
3328
3329
3330
3331
3332
3333
3334
3335
3336
3337
3338
3339
3340
3341
3342
3343
3344
3345
3346
3347
3348
3349
3350
3351
3352
3353
3354
3355
3356
3357
3358
3359
3360
3361
3362
3363
3364
3365
3366
3367
3368
3369
3370
3371
3372
3373
3374
3375
3376
3377
3378
3379
3380
3381
3382
3383
3384
3385
3386
3387
3388
3389
3390
3391
3392
3393
3394
3395
3396
3397
3398
3399
3400
3401
3402
3403
3404
3405
3406
3407
3408
3409
3410
3411
3412
3413
3414
3415
3416
3417
3418
3419
3420
3421
3422
3423
3424
3425
3426
3427
3428
3429
3430
3431
3432
3433
3434
3435
3436
3437
3438
3439
3440
3441
3442
3443
3444

```

```

602 602 GCTGAAAGAGCAAGAGGCTTTAAAGTACAGTCAAGAAATAGCTTTGATATTCAGAGTCA 721
QY 826 AspSerLeu 827
DB 122 CATGAA 727

RM454724 835 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6402743 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528814
5' mRNA sequence.
ACCESSION RM454724
VERSION RM454724.1 GI:18502764
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 835)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rquapbs@mail.nih.gov
Tissue Procurement: AFCC/DCIB/HP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Agencourt Bioscience Corporation (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: IAM2204 row: c column: 07
High quality sequence stop: 647.
Location/Qualifiers
1..835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528814"
/clone_1ib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="FHI10B (phage resistant)"
/vector="organ: skin; Vector: pCMV Sport6; Site: 1; NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Gligo di.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 254 a 210 c 137 g 243 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 8,056-110 Length: 835
Score: 1208.00 Matches: 244
Percent Similarity: 92.91% Conservatio: 5
Best Local Similarity: 91.04% Mismatches: 12
Query Match: 26.87% Indels: 7
Dns: 13 Gaps: 5

US 09 026 459A 49 (1 831) x RM454724 (1 835)

29 493 LeuAlaSerGlySerLeuHisAlaGluMetGluSerLeuAlaTrpLeuSerAspSerProLeu 512
33
DB 8 GTCGAGGAGAGAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 67

QY 513 PheAspLeuHisLeuGlyGlnSerIleGlyProThrAspHisLeuGlnSerAla 522
DB 68 TTGATCTATTATTAACCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 127

QY 533 CysTrpLeuAsnSerProLeuGlnSerProLeuGlnSerProLeuGlnSerProLeu 552
DB 128 TCTGCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 187

QY 554 ValAlaSerProLeuGlyGlySerProLeuGlySerProLeuGlySerProLeuGlySer 572
DB 188 GTAGAGATGTCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 247

```

```

QY 573 ThrGlnAlaThrSerAlaPheSerLeuGlnIleGlySerProLeuGlySerProLeuGlySer 592
DB 248 ACACAGAGCAAGCTCAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407

QY 593 PheTrpGlySerValIleTrpArgLeuAlaTrpIleAsnThrLeuGlySerGlnAlaGln 612
DB 308 TTCTATAAAAAAGTGATAGCTAGCTATCTCGGGTAAATATACCTTCTCAACCTGCTT 467

QY 613 LeuSerGlnHisProGlnLeuGlnHisIleLeuTrpThrLeuPheGlnHisThrLeuGln 632
DB 368 CTGTCTGAGCAAGCAAGAGATAGACATATATCTGACCTCTCTTCCACAGAGAGAGAGAG 427

QY 633 AsnGlnTrpGlnLeuGlnSerProLeuGlnSerProLeuGlnSerProLeuGlnSerPro 652
DB 438 AATGAGTATCAACTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 487

QY 653 GlyIleCysGlySerValGlySerProLeuGlnSerProLeuGlnSerProLeuGlnSer 672
DB 488 GTCATATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 547

QY 673 AspLeuProHisAlaValGlnGlnThrPheGlySerValGlnSerProLeuGlnSerPro 692
DB 548 GATCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607

QY 693 AspSerLeuHisAlaPheTrpAsnSerValPheMetGlnArgLeuGlySerProLeuGln 712
DB 608 GATCTATATAGTATTTCTATAATCTGCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 667

QY 713 GlnTrpAlaSerThrArgProPro---ThrLeuSerProLeuProHisHisProAspSer 731
DB 668 GAGTATGCTCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 727

QY 741 rProTrpTrpLys---PheProSerSerProLeu---AspGlnPro---GlyGlyAsn--- 746
DB 728 CCCCTTACAGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 787

QY 747 TleGlyTrpLeuSerProLeuLys 753
DB 788 TATATTTTAAAGCTCTGAGAA 809

RESULT 5
RM450031
LOCUS RM450031.1 GI:18459071
DEFINITION AGENCOURT_6493544 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528047
5' mRNA sequence.
ACCESSION RM450031
VERSION RM450031.1 GI:18459071
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1024)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rquapbs@mail.nih.gov
Tissue Procurement: AFCC/DCIB/HP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Agencourt Bioscience Corporation (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: IAM2204 row: d column: 22
High quality sequence stop: 643.
Location/Qualifiers
1..1024
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528047"

```







Tel: 314 266 1800  
 Fax: 314 266 1810  
 Email: mouse@atlas.wustl.edu  
 This clone is available royalty-free through LINC; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:661493  
 Seq primer: 28ml rev2 ET from Amersham  
 High quality sequence stop: 495.  
 Location/Qualifiers  
 1..871  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1247655"  
 /clone\_1fb="Soares mammary gland\_NMLMG"  
 /sex="female (lactating)"  
 /tissue\_type="mammary gland"  
 /lab\_host="DH10b"  
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified  
 polylinker; 1st strand cDNA was prepared from mammary  
 gland tissue from a lactating female, and was then primed  
 with a Not I - oligo(dT) primer. Double stranded cDNA was  
 ligated to Eco RI adaptors (Pharmacia), digested with Not  
 I and cloned into the Not I and Eco RI sites of the  
 modified pTZ19 vector. Library is normalized. Library  
 was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 242 a 172 c 204 g 251 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,416/90 Length: 871  
 Score: 1015.00 Matches: 216  
 Percent Similarity: 80.44% Conservative: 17  
 Best Local Similarity: 74.48% Mismatches: 53  
 Query Match: 22.53% Indels: 5  
 Gaps: 2

US 09 026 459a 49 (1 871) x AA761485 (1-871)  
 27 279 ThrTrpAspGlySerAsp...GCAAGGCTGluValAspValleProGProHistrProV 418  
 1b 860 AAGGCAAGAAAAACAGCGTTCGGAGCGGCAACGCGGTTCAGCGCCACAT---CCAG 804  
 27 418 aTArqThrValMetAsuThrIleGluGlnLeuMetMetIleLeuAsuSerAlaSerAspG 338  
 1b 803 TGGGCGGCTTCGAC 111  
 27 438 InProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProlylSerGluSerI 358  
 1b 744 GTCATCGAAGAAATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 111  
 27 458 IeLeuTysArqValLysAspIleLeuTyrIlePheLysSerIlePheLysSerIlePheLys 378  
 1b 684 TCCAGACGCTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 625  
 27 478 IyGlnIlyValValIleLeuLysSerIleLeuLysSerIleLeuLysSerIleLeuLysSerI 498  
 1b 624 GCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565  
 27 498 I4V4A-MetGlnSerMetLeuLysSerGluGluGluValLeuLeuSerIleGlnAsnProSerL 418  
 1b 664 GGTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 505  
 27 418 ystLeuLeuAspAspAspIlePhePheMetSerLeuLeuAlaLysAlaLeuGluValValIM 438  
 1b 604 CACCTGTAAGACAAAC 445  
 27 448 aTAlaThrTyrSerArqSerThrSerGlnAsuLeuAspSerGlyThrAspLeuSerPheP 458  
 1b 444 TGGTACGAC 385  
 27 458 TAlpIleLeuAsuValIleAsuLeuLysAlaPheAspPheTyrTysValIleGlnSerP 478  
 1b 404 GGTGATGCTGAAAGTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 325

478 heIleLysAlaGluGlyAsnLeuThrArqLeuMetIleLysHisLeuThrAlaGlySerGlu 498  
 1b 324 TTAICAAAGTGGAGGCACTTCACAGACAAAGATGAAAAATTTAAGAAATAGAGAC 265  
 498 isArqIleMetGluSerLeuAlaTyrLeuSerAspSerProLeuPheAspLeuIleLysG 518  
 1b 264 ATGAATCATGGAATGCTTGATGGCTTCAGATTACATTTATTTATTTATTTATTTAT 205  
 518 InSerLysAspArqGluGlyProIleAspHisLeuGlnSerAlaLysProLeuAsnLeu 538  
 1b 204 AGTCAAGAAAG 148  
 538 rGluGlnAsnAsnHisThrAlaAspMetTyrLeuSerProValAlaSerProLysL 558  
 1b 147 CTCGCCAGGGAACCAACTCCAGCAATGATGATGATGATGATGATGATGATGATGAT 88  
 558 yLysGlySerThrThrArqValAsuSerThrAlaAsuAlaGlnIleGlnIleGlnIle 578  
 1b 87 AAGAACATTCGCAATACAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 28  
 578 lApeGlnThrGlnLysProLeuLys 586  
 1b 27 CTTCCATATTCGAAAGCATTTGAAA 2

RESULT 10  
 HG253543 922 bp mRNA Lineat EST 14 FEB 2001  
 LOCUS 602364278F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:4472611 5'  
 DEFINITION mRNA sequence.  
 ACCESSION HG253543 GI:12763459  
 VERSION HG253543.1 GI:12763459  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cladacea; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 922)  
 AUTHORS NIH MGC http://www.ncbi.nlm.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: eugens@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA library preparation: Life Technologies, Inc.  
 cDNA library Arrayed by: the I.M.A.G.E. Consortium (LINC)  
 DNA Sequencing by: Inyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Platte, LAM10293 rev. 1 column, 20  
 High quality sequence stop: 633.  
 Location/Qualifiers  
 1..922  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4472611"  
 /clone\_1fb="NIH\_MGC\_90"  
 /tissue\_type="adipocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pCMV Spork6; Site 1: Not I;  
 Site 2: SalI; cloned unidirectionally; oligo dT primed.  
 Average insert size: 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC library."

BASE COUNT 307 a 214 c 204 g 197 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,060/89 Length: 922  
 Score: 1894.98 Matches: 224  
 Percent Similarity: 90.51% Conservative: 6  
 Best Local Similarity: 88.54% Mismatches: 21











Search completed: January 19, 2003, 05:07:26  
for time : 1939.01 secs

---



1. Carriacci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
557-573  
10449646

2. Carriacci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

3. Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carriacci, P., Konno, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Nagawa, Y., Izawa, M., Akara, F., Warahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIFEN integrated sequence analysis (PISA) system for full-length cDNA sequencing with 384 multi-capillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20540933  
11076861

4. Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arai, A., Hara, A., Fukuchi, Y., Kono, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, I., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gasterland, T., Gissi, C., King, K., Kochiwa, H., Kusl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasolo, C., Pasolo, M., Quackenbush, J., Schriml, M., Staib, P., Suzuki, P., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baidarrelli, R., Harsh, C., Blake, J., Hoffell, D., Hojunga, N., Carriacci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, M., Mazzarelli, J., Mombauts, F., Nordone, P., Kind, H., Kinwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyozaki, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 399 (6821), 685-690 (2001)  
11217851  
21085660

5. (bases 1 to 1364)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arai, T., Hara, A., Fukuchi, Y., Kono, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, I., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gasterland, T., Gissi, C., King, K., Kochiwa, H., Kusl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasolo, C., Pasolo, M., Quackenbush, J., Schriml, M., Staib, P., Suzuki, P., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baidarrelli, R., Harsh, C., Blake, J., Hoffell, D., Hojunga, N., Carriacci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, M., Mazzarelli, J., Mombauts, F., Nordone, P., Kind, H., Kinwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyozaki, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 399 (6821), 685-690 (2001)  
11217851  
21085660

6. (bases 1 to 1364)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arai, T., Hara, A., Fukuchi, Y., Kono, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, I., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gasterland, T., Gissi, C., King, K., Kochiwa, H., Kusl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasolo, C., Pasolo, M., Quackenbush, J., Schriml, M., Staib, P., Suzuki, P., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baidarrelli, R., Harsh, C., Blake, J., Hoffell, D., Hojunga, N., Carriacci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, M., Mazzarelli, J., Mombauts, F., Nordone, P., Kind, H., Kinwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyozaki, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 399 (6821), 685-690 (2001)  
11217851  
21085660

7. (bases 1 to 1364)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arai, T., Hara, A., Fukuchi, Y., Kono, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, I., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gasterland, T., Gissi, C., King, K., Kochiwa, H., Kusl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasolo, C., Pasolo, M., Quackenbush, J., Schriml, M., Staib, P., Suzuki, P., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baidarrelli, R., Harsh, C., Blake, J., Hoffell, D., Hojunga, N., Carriacci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, M., Mazzarelli, J., Mombauts, F., Nordone, P., Kind, H., Kinwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyozaki, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 399 (6821), 685-690 (2001)  
11217851  
21085660

8. (bases 1 to 1364)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arai, T., Hara, A., Fukuchi, Y., Kono, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, I., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gasterland, T., Gissi, C., King, K., Kochiwa, H., Kusl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasolo, C., Pasolo, M., Quackenbush, J., Schriml, M., Staib, P., Suzuki, P., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baidarrelli, R., Harsh, C., Blake, J., Hoffell, D., Hojunga, N., Carriacci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, M., Mazzarelli, J., Mombauts, F., Nordone, P., Kind, H., Kinwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyozaki, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 399 (6821), 685-690 (2001)  
11217851  
21085660

9. (bases 1 to 1364)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arai, T., Hara, A., Fukuchi, Y., Kono, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, I., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gasterland, T., Gissi, C., King, K., Kochiwa, H., Kusl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasolo, C., Pasolo, M., Quackenbush, J., Schriml, M., Staib, P., Suzuki, P., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baidarrelli, R., Harsh, C., Blake, J., Hoffell, D., Hojunga, N., Carriacci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, M., Mazzarelli, J., Mombauts, F., Nordone, P., Kind, H., Kinwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyozaki, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 399 (6821),









QY	745	-TCTTTTCScriptLeutys 751       :
DB	788	TATATTTCACCCCGGTGAA 809
RESULT 6		
BM450031		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
BASE COUNT		
ORIGIN		
Alignment Scores:		
Pred. No.:		
Score:		
Percent Similarity:		
Best Local Similarity:		
Query Match:		
DB:		
US-09-026-459A-47 (1-869) x BM450031 (1-1023)		
QY	654	LysValLysAsnIleAspSerLeuLysIleValThrAlaTyrlLysAspLeuPro 673
DB	2	AAATGAAGATAATAGACTTAATTCAAAATCATAATCATTAAGAGAGATGCT 61
QY	674	HisAlaValGlnGluThrPheLysArgValIleCylcLysGlnGlyArgSerIle 693
DB	62	CATGTCTTTTGAGAGCAATTCAGAGTGCTTTTATCAAGAAGAGAGAGATGATTTAT 121
QY	694	IleValPheTrpAsnSerValPheMetIleArAlaLeuLysThrAsnIleLeuAla 713
DB	122	ATAGTATCTATAACTGGCTCTTATACACAGATCAAAAAAATATTCACAGTAGCT 181
QY	714	SerThrArgProProThrLeuSerProIleProHisIleProArgSerProTrpLysPhe 743
DB	182	TCCACAAGAGCGGAGTACGTTGTCAGAAATATCAATATGATCAAGAGCTTCAAAGTT 241
QY	744	ProSerSerProLeuArgTrileProClValAsnIleValThrSerProLeuArgSer 753







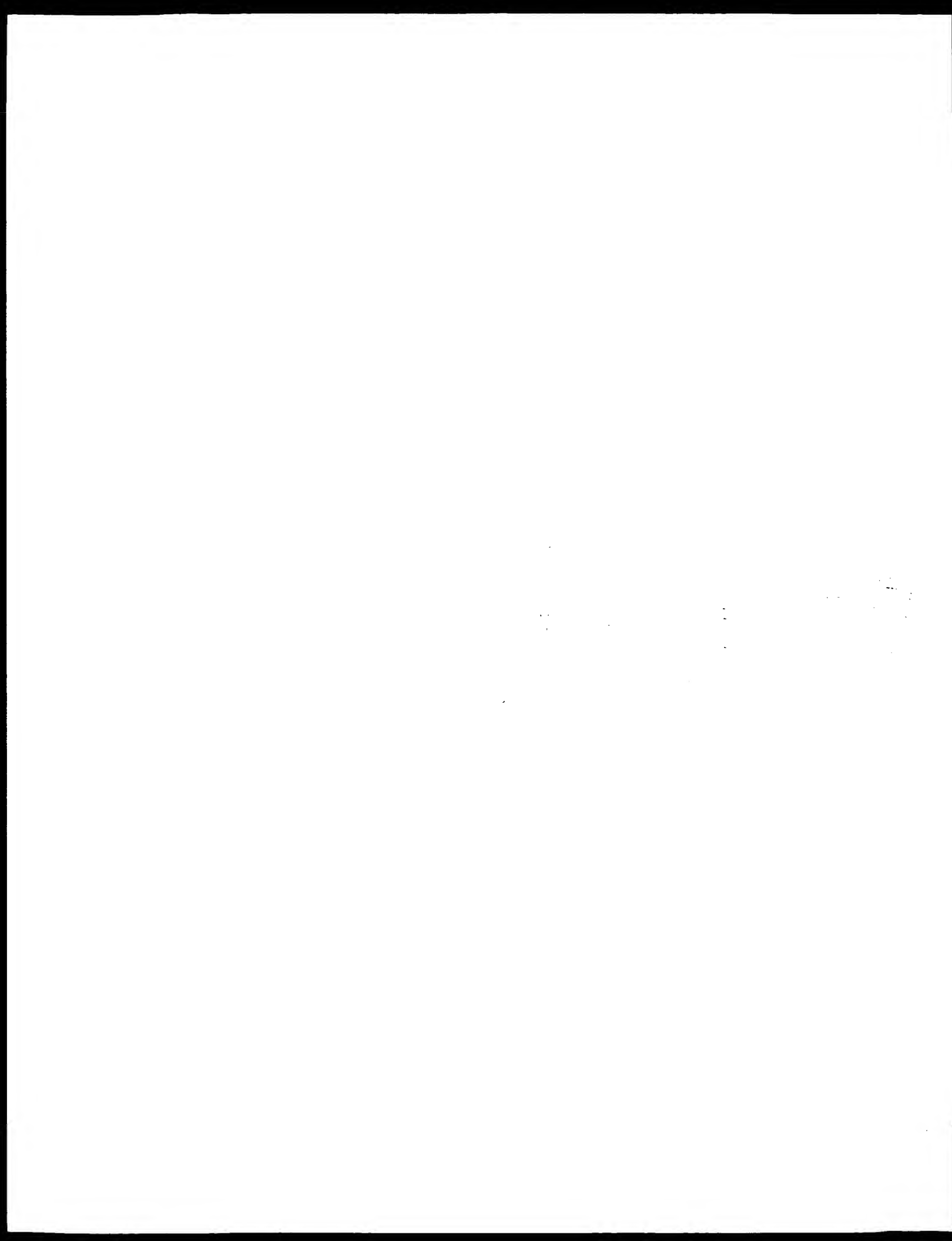












Copyright (c) 1993 - 2004, Compusero version 5.1.1.3

from multiple - nucleic acid search, using SW model

January 16, 2002, 15:20:22 ; Search time 00:01:31 Seconds  
(without alignments)  
18469, 506 Million cell updates/sec

US 09 026 459A 46

Perfect score: 337  
Sequence: | CCGCTCAGCGCGTTAAAAC.....AAATCAGGATTAATCATAG|| 337

Scoring table: IDENTITY\_NUM  
 0000 10 00 00000 1 0

See also: 16154(066) seeds, 809774 3376, residues

CH. 1000 c. 2. **State of Tenn. v. Robert Lee Smith** for a Requisition for Trial

Model	len	len2	len3	len4	len5	len6	len7	len8	len9	len10	len11	len12	len13	len14	len15	len16	len17	len18	len19	len20	len21	len22	len23	len24	len25	len26	len27	len28	len29	len30	len31	len32	len33	len34	len35	len36	len37	len38	len39	len40	len41	len42	len43	len44	len45	len46	len47	len48	len49	len50	len51	len52	len53	len54	len55	len56	len57	len58	len59	len60	len61	len62	len63	len64	len65	len66	len67	len68	len69	len70	len71	len72	len73	len74	len75	len76	len77	len78	len79	len80	len81	len82	len83	len84	len85	len86	len87	len88	len89	len90	len91	len92	len93	len94	len95	len96	len97	len98	len99	len100
Model	len	len2	len3	len4	len5	len6	len7	len8	len9	len10	len11	len12	len13	len14	len15	len16	len17	len18	len19	len20	len21	len22	len23	len24	len25	len26	len27	len28	len29	len30	len31	len32	len33	len34	len35	len36	len37	len38	len39	len40	len41	len42	len43	len44	len45	len46	len47	len48	len49	len50	len51	len52	len53	len54	len55	len56	len57	len58	len59	len60	len61	len62	len63	len64	len65	len66	len67	len68	len69	len70	len71	len72	len73	len74	len75	len76	len77	len78	len79	len80	len81	len82	len83	len84	len85	len86	len87	len88	len89	len90	len91	len92	len93	len94	len95	len96	len97	len98	len99	len100

Test processed: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

**Abstract**

1: em\_est\_b0: \*  
2: em\_est\_bum: \*  
3: em\_est\_l0: \*  
4: em\_est\_l1: \*  
5: em\_est\_mu: \*  
6: em\_est\_ov: \*  
7: em\_est\_pl: \*  
8: em\_est\_po: \*  
9: em\_est\_s: \*  
10: qb\_est\_1: \*  
11: qb\_est\_2: \*  
12: qb\_est\_3: \*  
13: qb\_est\_4: \*  
14: qb\_est\_5: \*  
15: qb\_est\_l0: \*  
16: qb\_est\_mu: \*  
17: qb\_est\_ov: \*  
18: qb\_est\_pl: \*  
19: qb\_est\_po: \*  
20: qb\_est\_s: \*  
21: qb\_qss\_bum: \*  
22: qb\_qss\_l0: \*  
23: qb\_qss\_l1: \*  
24: qb\_qss\_mu: \*  
25: qb\_qss\_ov: \*  
26: qb\_qss\_pl: \*  
27: qb\_qss\_po: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	910.2	27.3	1023	BM450031	BM450031 AGENCOURT
2	909.24	26.9	1051	BM651083	BM651083 AGENCOURT
3	896.6	25.3	1081	BQ439221	BQ439221 AGENCOURT
4	792.4	22.3	1122	BM456881	BM456881 AGENCOURT
5	743.8	21.9	881	BQ220375	BQ220375 AGENCOURT
6	725.4	21.5	727	A1547813	A1547813 DKFZp3130

7	699.6	20.7	92.3	12	RC6.184.1
8	692.2	20.5	76.7	12	RC5.738.5
9	691.8	20.5	97.2	13	BM4.661.94
c	691.2	20.5	79.2	9	AL.56.3757
10	678	20.1	83.5	13	BM4.5.4724
11	660	19.5	67.2	13	RI868.59
12	660	19.5	130.1	13	AK01.1246
13	646	19.1	73.7	10	RC5.832.78
14	643.6	19.1	68.1	12	RC6.16.1.19
15	642.2	19.0	68.1	13	BM2.640.99
c	616	17.9	61.8	13	BM2.640.99
16	617	18.3	61.8	13	BM2.640.99
17	614	18.2	69.4	10	RC08.284.6
c	605.2	17.9	62.3	10	AW58.3181
18	605.2	17.9	60.5	9	AL.5996.33
19	605	17.9	92.2	12	RC2.535.43
20	596.4	17.7	92.2	12	RC2.535.43
21	580	17.2	58.0	9	AL.5991.05
c	575.2	17.0	91.4	9	AA76.44.1
22	566.4	16.8	58.8	10	AA76.44.1
c	558	16.5	87.1	9	AA76.348.5
24	540.2	15.7	53.8	10	AW58.2967
25	530	15.7	53.8	9	AL.69.2790
26	530	15.7	53.8	9	AL.69.2790
27	528	15.6	57.2	5	AA2.209.41
c	516.8	15.3	59.7	12	RC9.312.46
28	516.8	15.3	59.7	10	AA71.55.33
29	513.2	15.2	59.4	10	AA71.55.33
30	512.2	15.2	73.4	12	RC.34.8.57
31	508.6	15.1	115.4	14	RC80.27.36
32	505.2	15.0	73.6	13	RI1.5156.8
33	493.8	14.6	52.5	9	AL.58.25.5
34	486.4	14.4	58.4	12	RC14.905.0
35	484.4	14.3	43.3	10	AW50.2887
36	484.2	14.3	100.1	13	RC47.560.3
37	481.4	14.3	48.4	10	RI1.680.95
38	480.2	14.2	65.9	13	RI1.211.16
39	476.6	14.1	208.3	13	BM4.567.28
40	474.4	14.0	73.4	13	RI.5.26.98.2
41	465.6	13.8	88.0	14	RC2.22.22.7
42	459.2	13.6	49.7	12	RC9.34784
c	457	13.5	45.7	9	AL.09.32.15
43	457	13.5	73.2	9	AL.598766
44	455.4	13.5	64.3	13	RI.5.53.93.7
45	455.4	13.5	64.3	13	RI.5.53.93.7
46	452.4	13.4	64.3	13	RI.5.53.93.7
47	452.4	13.4	64.3	13	RI.5.53.93.7
48	452.4	13.4	64.3	13	RI.5.53.93.7
49	452.4	13.4	64.3	13	RI.5.53.93.7
50	452.4	13.4	64.3	13	RI.5.53.93.7
51	452.4	13.4	64.3	13	RI.5.53.93.7
52	452.4	13.4	64.3	13	RI.5.53.93.7
53	452.4	13.4	64.3	13	RI.5.53.93.7
54	452.4	13.4	64.3	13	RI.5.53.93.7
55	452.4	13.4	64.3	13	RI.5.53.93.7
56	452.4	13.4	64.3	13	RI.5.53.93.7
57	452.4	13.4	64.3	13	RI.5.53.93.7
58	452.4	13.4	64.3	13	RI.5.53.93.7
59	452.4	13.4	64.3	13	RI.5.53.93.7
60	452.4	13.4	64.3	13	RI.5.53.93.7
61	452.4	13.4	64.3	13	RI.5.53.93.7
62	452.4	13.4	64.3	13	RI.5.53.93.7
63	452.4	13.4	64.3	13	RI.5.53.93.7
64	452.4	13.4	64.3	13	RI.5.53.93.7
65	452.4	13.4	64.3	13	RI.5.53.93.7
66	452.4	13.4	64.3	13	RI.5.53.93.7
67	452.4	13.4	64.3	13	RI.5.53.93.7
68	452.4				

## ALIGNMENT

```

RESULT 1
BM450031
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
EST.
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:552837"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-Sport6, Site:1 NotI, Site:2: SalI; cloned unidirectionally. Primer: oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
BASE COUNT      344 a      211 c      197 g      281 t
ORIGIN

Query Match      27.0%; Score 910.2; DB 13; Length 1023;
Best Local Similarity 95.2%; Pred. No. 6.7e 168;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

CY 1965 CAAAGTGAAGAAATATAGCTTAAATTCAAATCATCTGTAAAGATACAGAAAGATCTCC 2024
DB 1 CAAAGTGAAGAAATATAGCTTAAATTCAAATCATCTGTAAAGATACAGAAAGATCTCC 60
CY 2025 TCATGCTGTTTCAGACAGATATCAAAAGCTGTTCATCAAGAGACAGACATCATCTAT 2084
DB 61 TCATGCTGTTTCAGACAGATATCAAAAGCTGTTCATCAAGAGAGAGATGATCTAT 120
CY 2085 TATAGTATTCATATACCTGCTCTCATGACAGACAGAAACAAATATTTTCAGATATGC 2144
DB TATAGTATTCATATACCTGCTCTCATGACAGACAGAAACAAATATTTTCAGATATGC 180
CY 2145 TTCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2204
DB TTCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2204
CY 2205 TTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 240
DB TTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 240
CY 2405 TCCATATATATATATATATATATATATATATATATATATATATATATATATATAT 2264
DB TCCATATATATATATATATATATATATATATATATATATATATATATATATATAT 300
CY 2265 ATATATAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2324
DB ATATATAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
CY 2325 ATATATAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2384
DB ATATATAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
CY 2385 ATGTAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2444
DB ATGTAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
CY 2445 GAAAAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2504
DB GAAAAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
CY 2505 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2564
DB AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
CY 2565 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2624
DB AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
CY 2625 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2684
DB GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
CY 2685 TAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2744
DB TAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 780
CY 2745 AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2802
DB AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
CY 2803 GTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2862
DB GTATATATATATATATATATATATATATATATATATATATATATATATATATAT 900
CY 2863 TTCTTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2918
DB TTCTTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
CY 2919 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2978
DB ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1020
CY 2979 A 2979
DB 1021 A 1021

RESULT 2
BM561083
DEFINITION      BM561083 1051 bp mRNA linear EST 20-FEB-2002
                    AGENCOURT.6566133 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550905
                    5' mRNA sequence.
ACCESSION      BM561083
VERSION        BM561083.1 GI:18806043
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mollusca; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1051)
                NIH-MGC http://mgi.nhlbi.nih.gov/.
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cga@bbs.nih.gov
                Tissue procurement: ATCC
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMRL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNI at:
                http://image.llni.gov
                Plate: LLML2263 row: m column: 18
                High quality sequence step: 737.
FEATURES
    source
        location/Qualifiers
            1..1051
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5550905"
                /clone_lib="NIH_MGC_67"
                /tissue_type="retinoblastoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: eye; Vector: pCMV-Sport6; Site:1: NotI;
                Site:2: SalI; Cloned unidirectionally. Primer: oligo dT.
                Average insert size 1.75 kb. Library constructed by Life
                Technologies."
BASE COUNT      330 a      221 c      197 g      299 t
ORIGIN

Query Match      26.9%; Score 909.4; DB 13; Length 1051;
Best Local Similarity 98.4%; Pred. No. 9.6e 168;
Matches 946; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

CY 1123 AAAATCTCTGCAATATATATATATATATATATATATATATATATATATATATAT 1182
DB 1 AAAATCTCTGCAATATATATATATATATATATATATATATATATATATATATAT 60
CY 1183 TTCTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1242
DB TTCTATATATATATATATATATATATATATATATATATATATATATATATATATAT 120
CY 1243 AATTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1302
DB AATTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1302

```



us-09-026-459a-46.rst

Tue Jan 21 10:04:22 2003

```

3199 TAGAAATAGAAAAAATTACTAATTTTACACATTAGATTTTATTTTACTATGGAATCT 3198
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
3200 CTCAAAAGAAATGCTGAAGGAGCAACCTCTTAACCACTGAAAAAATGCTTTGAT 359
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2464 ATTGAAGGATCAGATGAAGAGCAGATGGAAGTAACATCTCCAGCAGATCCAAATTCAG 2523
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
360 ATTGAAGGATCAGATGAAGAGCAGATGGAAGTAACATCTCCAGCAGATCCAAATTCAG 419
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2524 CAGAAATCTGGAGAAATGACTTCTACTCGAACAGGAATGAGGATCTCAGGACCTTGTGACACATGT 2583
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
420 CAGAAATCTGGAGAAATGACTTCTACTCGAACAGGAATGAGGATCTCAGGACCTTGTGACACATGT 479
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2584 AGCATGATACCTTCAACAAAGAGAGAAATGAGGATCTCAGGACCTTGTGACACATGT 2643
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
480 AGCATGATACCTTCAACAAAGAGAGAAATGAGGATCTCAGGACCTTGTGACACATGT 539
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2644 GTACACCTCTGGATTCATTTCTCTCACAGATGCTGATTAACATTTCCAGGCTTCTG 2703
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
540 GTACACCTCTGGATTCATTTCTCTCACAGATGCTGATTAACATTTCCAGGCTTCTG 599
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2704 TTATGGCCACATTTAATATCTTTCAGCTCTTTTGTGGATATAAAATGTCAGATGCAAT 2763
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
600 TTATGGCCACATTTAATATCTTTCAGCTCTTTTGTGGATATAAAATGTCAGATGCAAT 659
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2764 GTTGGGTGATTCCTTAAGCCACCTTGAATGTTAGTCAITGTTATTATACAGATTGAAA 2823
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
660 GTTGGGTGATTCCTTAAGCCACCTTGAATGTTAGTCAITGTTATTATACAGATTGAAA 719
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2824 ATCTGTGTAATCTCTGCAATTTAAAGTGTAGCAGATGTTTCTCTTCCCAAGIAA 2883
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
720 ATCTGTGTAATCTCTGCAATTTAAAGTGTAGCAGATGTTTCTCTTCCCAAGIAA 779
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2884 ATTTGTTTTCATTTATGGATAGTAAATGTCCTTAAGTGGAGTCTGATAGCCAG 2943
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
780 AATTGGTGTCTTTATGGATAGTAAATGTCCTTAAGTGGAGTCTGATAGCCAG 839
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
2944 CTTGCTGACTTCTTCTTCTTTTGTAGCATATAGTGTGTTTCTCTTCTTCTTAT 898
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
840 CTTGCTGACTTCTTCTTCTTTTGTAGCATATAGTGTGTTTCTCTTCTTCTTAT 898
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
3004 TAATTTATA-TCTAIAI|||||TAATTTACACAA-CACTTACAGAAATGTCCTA 3060
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
899 TAATTTATA-TCTAIAI|||||TAATTTACACAA-CACTTACAGAAATGTCCTA 958
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
3061 TCTATCTTCCAAATGCAATTTGATGCTGCTTACCAAAATTTATCTGTAAC-CTC 3120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
959 TCTATCTTCCAAATGCAATTTGATGCTGCTTACCAAAATTTATCTGTAAC-CTC 1017
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
3121 TGCAGAAAT-GGATATTATTAGAAATAGAAAAAATTTACTAATTTTACACATTAGATT 3179
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1018 TGCAGAAATGGGAATATTATTAGAAATAGAAAAAATTTACTAATTTTCCCTTAAAT 1077
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
3180 TATTTTACTATTGCAATCTGATATA 3204
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1078 TATTTTACTATTGCAATCTGATATA 1102
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 5
HQ220275 881 bp mRNA linear EST 02-MAY-2002
Locus AGENCOURT_7572621 NIH_MGC_92 Homo sapiens cDNA clone IMAGE 588245
DEFINITION 5', mRNA sequence.
VERSION HQ220275
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
ACCESSION NIH-MGC http://mgi.ncbi.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

```

```

RESULT 4
BM545881 1122 bp mRNA linear EST 20-FEB-2002
Locus AGENCOURT_6505183 NIH_MGC_125 Homo sapiens cDNA clone IMAGE 588245
DEFINITION 5', mRNA sequence.

```

```

ACCESSION BM545881
VERSION BM545881
KEYWORDS EST
SOURCE EST

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS NIH-MGC http://mgi.ncbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cnapts@mail.nih.gov

```

```

Tissue Procurement: Invitrogen
cDNA Library Prepared by: Life Technologies, Inc.
cDNA Library Arrayed by: The T.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the T.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12359, row: a, column: 14
High quality sequence start: 48
High quality sequence stop: 773.
Location/Qualifiers
1..1122 "Homo sapiens"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:588245"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site 1: EcoRV (destroyed); Site 2: NotI; PNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1.3-5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Griber (Invitrogen). Research Genetics
tracking code 036."
360 a 185 g 359 t 2 others
BASE COUNT 360 a 185 g 359 t 2 others

```

## FEATURES

## SOURCE

```

Query Match 22.3% Score 752.1 Pw 13. Length 1122:
Best Local Similarity 95.6% Pw 5e-137:
Matches 827: Conservative 0: Mismatches 32: Indels 6: Gaps 5:

```

```

QY 2344 TTGGCCATCTTGAGAGCTTCCAGAAATTAATACAGATGCTATGTAACAGGATGCTG 2403
||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 240 TTGACATCTTCTCAAGACTTCCAGAAATTAATACAGATGCTATGTAACAGGATGCTG 299
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2404 CTCAAAAGAGTCTGAAGAGCAACCTCTTAACCACTGAAAAAATGCTTTGAT 2463
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```















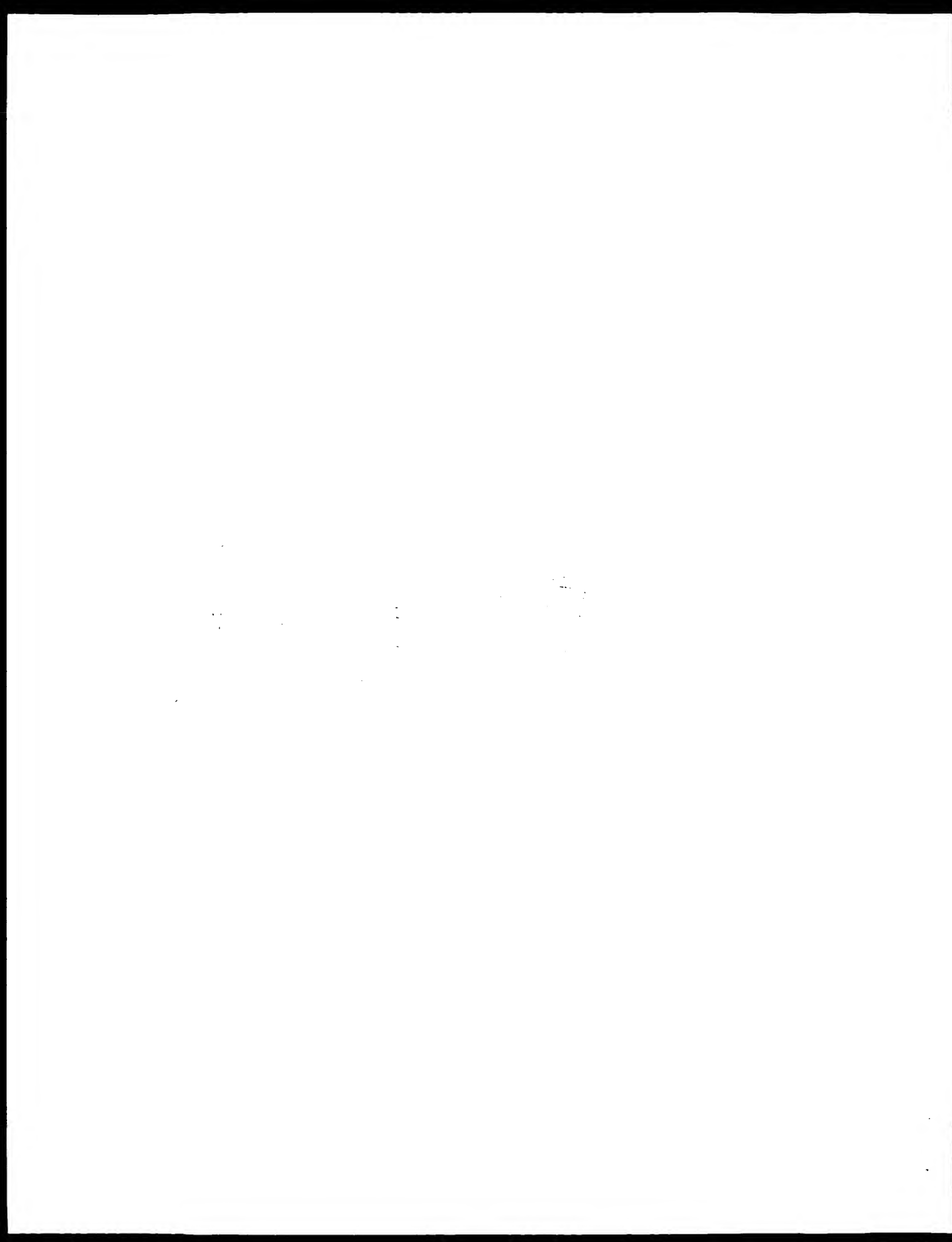


```

Query Match      19.08; Score 642.2; DB 12; Length 681;
Best Local Similarity 98.48; Pred. No. 1.9e-115;

```







Genome version 5.1.4  
Copyright (c) 1993 - 2003 Computer Ind.

EM nucleic nucleic search using SW model

Run on: January 16, 2003, 15:20:22 : Search time 2974.47 seconds  
(without alignments)  
18469.506 Million cell updates/sec

Filter: US 09 026-459A to  
Perfect score: 3392  
Sequence: 1 GGCATGATGATTTTACTG.....AAATGAGATTTATGATGACT 3392

Scoring table: IDENTITY 100  
Gapop 10.0, Gapext 1.0

Scanned: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42406142

Minimum hit seq length: 0  
Maximum hit seq length: 2000000000

Post processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
EST :			
1:	em. est1.1	13	em. est1.1
2:	em. est1.2	21	em. est1.2
3:	em. est1.3	31	em. est1.3
4:	em. est1.4	41	em. est1.4
5:	em. est1.5	51	em. est1.5
6:	em. est1.6	61	em. est1.6
7:	em. est1.7	71	em. est1.7
8:	em. est1.8	81	em. est1.8
9:	em. est1.9	91	em. est1.9
10:	em. est1.10	101	em. est1.10
11:	em. est1.11	111	em. est1.11
12:	em. est1.12	121	em. est1.12
13:	em. est1.13	131	em. est1.13
14:	em. est1.14	141	em. est1.14
15:	em. est1.15	151	em. est1.15
16:	em. est1.16	161	em. est1.16
17:	em. est1.17	171	em. est1.17
18:	em. est1.18	181	em. est1.18
19:	em. est1.19	191	em. est1.19
20:	em. est1.20	201	em. est1.20
21:	em. est1.21	211	em. est1.21
22:	em. est1.22	221	em. est1.22
23:	em. est1.23	231	em. est1.23
24:	em. est1.24	241	em. est1.24
25:	em. est1.25	251	em. est1.25
26:	em. est1.26	261	em. est1.26
27:	em. est1.27	271	em. est1.27

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	910.2	26.8	1024	13	BM450031	BM450031 AGNCOURT
2	909.4	26.8	1051	13	BM561084	BM561084 AGNCOURT
3	879.8	25.9	1364	11	AK011246	AK011246 Mus muscu
4	856	26.2	1081	14	BQ439221	BQ439221 AGNCOURT
5	844.8	24.9	881	14	BQ429275	BQ429275 AGNCOURT
6	786	24.2	1001	13	BM475603	BM475603 AGNCOURT

7	752.4	22.2	1122	13	BM450031	BM450031 AGNCOURT
8	725.4	21.4	727	9	AL597811	AL597811 DKFZP43146
9	720.6	21.2	896	14	BQ222227	BQ222227 AGNCOURT
10	699.6	20.6	824	12	BG010061	BG010061 60261848
11	698.2	20.6	792	9	AL563757	AL563757 AL563757
12	692.2	20.4	747	12	BQ573156	BQ573156 G62591412
13	691.8	20.4	972	13	BM466194	BM466194 AGNCOURT
14	678	20.0	835	13	BM454724	BM454724 AGNCOURT
15	660	19.5	674	13	B1868459	B1868459 60492450
16	643.6	19.0	797	10	BE539278	BE539278 601065567
17	642.2	18.9	581	12	BG614219	BG614219 602644717
18	617	18.2	618	13	BM264009	BM264009 1944408.X
19	614	18.1	694	10	BE082846	BE082846 KC2 H1064
20	605.2	17.8	623	10	AW583181	AW583181 1311309.Y
21	605	17.8	605	9	AL599643	AL599643 DKFZP43146
22	596.4	17.6	922	12	BQ253543	BQ253543 602644278
23	580	17.1	580	9	AL599105	AL599105 DKFZP43146
24	575.2	17.0	914	9	AA763411	AA763411 VW54304.1
25	566.4	16.7	588	10	AW368234	AW368234 CM3 H1018
26	558	16.5	871	9	AA763485	AA763485 VW54101.1
27	539.2	15.6	598	10	AW582567	AW582567 1A0608.Y
28	530	15.6	549	9	AL692790	AL692790 DKFZP43146
29	528	15.6	572	9	AL120941	AL120941 DKFZP43146
30	516.8	15.2	597	12	HE541236	HE541236 MR0 H1015
31	513.2	15.1	594	10	AV715533	AV715533 AV715533
32	512.2	15.1	734	12	BF144857	BF144857 60179044.4
33	505.2	14.9	736	13	BF151568	BF151568 602916108
34	502.4	14.8	832	9	AL646048	AL646048 VW54318.X
35	493.8	14.6	525	9	AA258255	AA258255 2159404.1
36	486.4	14.3	584	12	HG149050	HG149050 udbrc01.Y
37	484.4	14.3	493	10	AW502887	AW502887 U1-HF HN0
38	481.4	14.2	484	10	HE168095	HE168095 QV3 H1051
39	480.2	14.2	659	13	BF182116	BF182116 604035591
40	476.6	14.1	2083	13	BM456728	BM456728 AGNCOURT
41	474.4	14.0	734	13	B1526982	B1526982 602927377
42	459.3	13.5	497	12	BF934784	BF934784 KC3 H1094
43	457	13.5	457	9	AL093215	AL093215 qd0402.X
44	457	13.5	742	9	AL598766	AL598766 DKFZP43146
45	454.2	13.4	869	12	BE865749	BE865749 601678131

ALIGNMENTS

RESULT 1  
BM450031  
LOCUS AGNCOURT\_6493544 NIH\_MGC\_72 Homo Sapiens cDNA clone IMAGE:5528037  
DEFINITION 1024 bp mRNA linear EST 05-FEB-2002  
5' mRNA sequence  
ACCESSION BM450031 GI:18499071  
VERSION BM450031.1  
KEYWORDS EST  
SOURCE human  
ORGANISM Homo Sapiens

REFERENCE 1 (bases 1 to 1023)  
AUTHORS NIH-MGC <http://mdc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [graphs-r@mail.nih.gov](mailto:graphs-r@mail.nih.gov)  
Tissue Procurement: ATCC/DCTD/UTP

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: CLAM12204 row 3 column 22  
high quality sequence stop: 643  
Location/Qualifiers 1..1023







http://image.llnl.gov

Plate: LLAM3287 Row: k Column: 01

High quality sequence stop: 732

## FEATURES

Source Location/Qualifiers

1..1081

Organism: Homo sapiens

db\_xref: taxon:9606

clone: IMAGE:6015480

clone: NIH\_MGC\_92

Tissue type: embryonal

Lab host: DH10B (phage-resistant)

Note: Organ: testis; Vector: pCMV Sport6; Site: 1; Note:

Site: 2; Salt: Cloned unidirectionally; oligo-dt primed.

Average insert size 2.5 kb. Library enriched for

full length clones and constructed by Life Technologies

Note: This is a NIH\_MGC library.

BASE COUNT 337 a 183 g 371 t

## ORIGIN

Query Match 25.2% Score 856; DB 14; Length 1081;

Best Local Similarity 99.08; Pred. No. 6; 155;

Matches 892; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

27 2494 GAAAGATGAGGAAATAAATATGTCGACGAGAGTCCAAATTCAGAGAACTGGAGAA 2553

DB 1 GAAAGATGAGGAAATAAATATGTCGACGAGAGTCCAAATTCAGAGAACTGGAGAA 60

27 2594 ATGATTTCTATCTGAAATGCAATGCAAAAGCAAAATGAAATGATGATGATGATGAT 2613

DB 61 ATGATTTCTATCTGAAATGCAATGCAAAAGCAAAATGAAATGATGATGATGATGAT 120

27 2614 AACAGAGAGAGAGAGAGAGATCTAGAGATCTGATGAGATCTGATGAGATCTGATG 2673

DB 121 AACAGAGAGAGAGAGAGATCTAGAGATCTGATGAGATCTGATGAGATCTGATG 180

27 2674 CATCTCTCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2733

DB 181 CATCTCTCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240

27 2734 ATATCTCTCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2793

DB 241 ATATCTCTCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400

27 2794 AAGGATTTGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2853

DB 401 AAGGATTTGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360

27 2854 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2913

DB 461 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

27 2914 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2973

DB 421 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

27 2974 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3033

DB 481 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

27 3034 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3093

DB 541 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

27 3094 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3153

DB 601 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

27 3154 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3213

DB 661 TGGATTTTAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719

27 3214 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3273

DB 181 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240

DB 720 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778

QY 3274 AGTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3334

DB 774 AGTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838

QY 3334 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3392

DB 834 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898

RESULT 5

QY 3392 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3452

DB 3452 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4092

QY 4092 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4152

DB 4152 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4792

QY 4792 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4852

DB 4852 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5492

QY 5492 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5552

DB 5552 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6192

QY 6192 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6252

DB 6252 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6892

QY 6892 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6952

DB 6952 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7592

QY 7592 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7652

DB 7652 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8292

QY 8292 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8352

DB 8352 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8992

QY 8992 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9052

DB 9052 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9692

QY 9692 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9752

DB 9752 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10392

QY 10392 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10452

DB 10452 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11092

QY 11092 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11152

DB 11152 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11792

QY 11792 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11852

DB 11852 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12492

QY 12492 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12552

DB 12552 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13192

QY 13192 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13252

DB 13252 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13892

QY 13892 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13952

DB 13952 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14592

QY 14592 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14652

DB 14652 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15292

QY 15292 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15352

DB 15352 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15992

QY 15992 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16052

DB 16052 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16692

QY 16692 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16752

DB 16752 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17392

QY 17392 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17452

DB 17452 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18092

QY 18092 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18152

DB 18152 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18792

QY 18792 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18852

DB 18852 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 19492

QY 19492 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 19552

DB 19552 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20192

QY 20192 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20252

DB 20252 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20892

QY 20892 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20952

DB 20952 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 21592

QY 21592 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 21652

DB 21652 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22292

QY 22292 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22352

DB 22352 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22992

QY 22992 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23052

DB 23052 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23692

QY 23692 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23752

DB 23752 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24392



QY 1041 GTTATGAAATTTGATTT GTATTATTAAGTGTGACATGTAATGCAAAAGAAAGATAT 1088  
 DB 824 CTCTAAAAAATCGAATTTCTTTTATTTTAACTGACGCTGCAATTCCTAAAGAAAGC 862

QY 1089 ATTGAAGACAGAGATAGAGATAGATGATGATGATGATGATGATGATGATGATGAT 1124  
 DB 884 TATATGGAAGAAAGAGAGATAGATGATGATGATGATGATGATGATGATGATGAT 918

RESULT 7  
 LOCUS BM545881 1122 bp mRNA Linear EST 20 FEB 2002  
 DEFINITION AGENE08RT.5505184 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5588245  
 5', mRNA sequence.  
 ACCESSION BM545881  
 VERSION 1  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
 1 (bases 1 to 1122)  
 REFERENCE NIH-MGC <http://mimc.nhlbi.nih.gov/>  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cepbes@mit.edu](mailto:cepbes@mit.edu)  
 Tissue Procurement: Invitrogen  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Applied Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/NIH at:  
<http://image.llnl.gov>  
 plate: L1AM1259 Row: a Column: 14  
 High quality sequence start: 48  
 High quality sequence stop: 773  
 FEATURES Location/Qualifiers  
 Source 1..1122  
 Zonedism="Homo sapiens"  
 Zdb-ref="taxon:9606"  
 Zclone="IMAGE:5588245"  
 Zclone.lib="NIH\_MGC\_125"  
 Zlab-host="DB105"  
 Note: "Ovarian ovary (pool of 3); Vector: pCMV-Sport16;  
 Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool  
 of three ovaries, from females ranging in age from 38 to  
 49 yo. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert,  
 size 2.1 kb, insert size range 1-3.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Graber (Invitrogen), Research Genetics  
 tracking code 036."
 BASE COUNT 460 a 185 c 216 g 459 t 2 others  
 ORIGIN

Query Match  
 Best Local Similarity 75.2% Score 752.4; DB 13; Length 1122;  
 Matches 827; Conservative 0; Mismatches 42; Indels 6; Gaps 5;

QY 2459 TTCTGGATTTTGAATTCAGAAAATAATCAGATGATGATGATGATGATGATGATGATGAT 2418  
 DB 240 TTTGACATTTGACAAATTCACAAAATAATTCAGATGATGATGATGATGATGATGATGAT 209

QY 2419 CTCTAAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2478  
 DB 400 CTCTAAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359

QY 2479 ATTGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2538  
 DB 460 ATTGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419

QY 2539 CATAAATGCTGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2598

DB 420 CATAAATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479  
 QY 2599 AGCATGATACCTGCAAAACAGCAGCAGCAAAACAGCAGCAGCAGCAGCAGCAGCAGCAG 2658  
 DB 480 AGCATGATACCTGCAAAACAGCAGCAGCAAAACAGCAGCAGCAGCAGCAGCAGCAGCAG 549  
 QY 2659 GTACAGCTCTGATTCATGCTCTCTCAATGATGATGATGATGATGATGATGATGATGAT 2718  
 DB 540 GTACAGCTCTGATTCATGCTCTCTCAATGATGATGATGATGATGATGATGATGATGAT 599  
 QY 2719 TTATGGCAGCATTTAATTAATCTTCTGCTCTTCTTCTGATATATAAATGATGATGAT 2778  
 DB 600 TTATGGCAGCATTTAATTAATCTTCTGCTCTTCTTCTGATATATAAATGATGATGAT 659  
 QY 2779 GTTGGGTCATTCCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2838  
 DB 660 GTTGGGTCATTCCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719  
 QY 2839 ATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2898  
 DB 720 ATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779  
 QY 2899 AATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2958  
 DB 780 AATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839  
 QY 2959 CTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3018  
 DB 840 CTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898  
 QY 3019 TAATTATA--TGATATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 3075  
 DB 860 TAATTATAATGGAATATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 918  
 QY 3076 TGTATCTTCCAAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3135  
 DB 959 TGTATCTTCCAAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017  
 QY 3136 TGTCAAAAT--GATATATTTAGCAATTTAGCAATTTAGCAATTTAGCAATTTAGCAATTT 3194  
 DB 1018 TGTCAAAATGCAATTTATTTAGCAATTTATTTAGCAATTTATTTAGCAATTTATTTAG 1077  
 QY 3195 TATTTATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3219  
 DB 1078 TATTTATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1102

RESULT 8  
 LOCUS AL597811  
 DEFINITION DKEZp41300816\_r1 413 (synonym: hlee2) Homo sapiens cDNA clone  
 DKEZp41300816 5', mRNA sequence.  
 ACCESSION AL597811  
 VERSION 1  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;  
 1 (bases 1 to 727)  
 REFERENCE Poustka A., Wellenreuther, R., Mewes, H.W., Well, H. and Wiemann, S. (1999)  
 EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Well, H. and Wiemann, S.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Poustka A.J.  
 Department Leirach  
 Max Planck Institute for Molecular Genetics  
 Ihnestr. 74, 14195 Berlin, Germany  
 Tel: +49-30-84131623  
 Fax: +49-30-84131128  
 Email: [poustka@mpg-berlin.dahlem.mpg.de](mailto:poustka@mpg-berlin.dahlem.mpg.de)  
 This is the 5' sequence of the clone insert





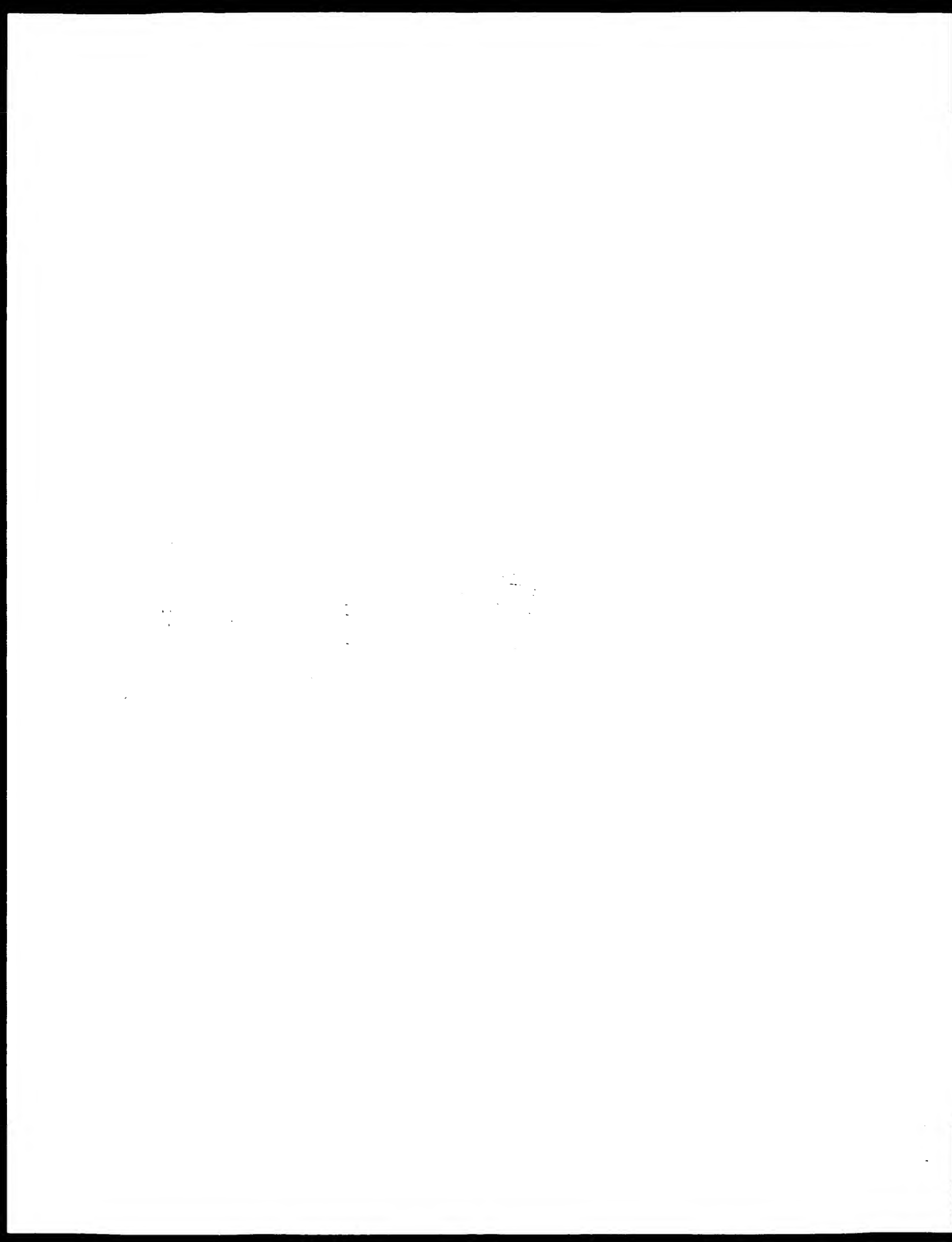


















```

QY 82 TysGluThrAspThrSerThrTyrValAspAsnAlaMetSerArgGluGluGlyTyr 101
DB 548 AAGAAATGATACACAGTACAGCTTATGATGCTATGCTAGAGACTATTGAAGAGATAT 597
QY 102 AspValLeuPheAlaLeuPheSerTyrLeuGluArgThrCysGluLeuLeuTyrLeuThr 121
DB 548 AATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
QY 122 GluProSerSerSerThrGluThrGluThrLeuValLeuValLeuValSerTyr 141
DB 548 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 142 ThrThrPheGluLeuAlaTyrGlyGluValLeuGluMetGluAspGluValThrSer 161
DB 718 ATGCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
QY 162 PheGluLeuMetLeuGlyValLeuAspThrPheThrLeuLeuSerProMetLeuLeu 181
DB 778 ATGCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847
QY 182 TysGluProTyrTyrThrAlaValThrProLeuAsnGlySerProArgThrProArgArg 201
DB 848 AGCAAGCTCAACAAAAGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
QY 202 GlyGluAsnArgSerAlaArgThrAlaTyrGluLeuGluAsnAspThrArgThrLeuGlu 221
DB 898 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
QY 222 ValLeuGlyTyrGluThrSerGluValLeuAspGluValLeuValLeuValLeuVal 241
DB 958 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
QY 242 PheThrProMetAspSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGlu 261
DB 1018 ATGCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
QY 262 AsnLeuSerTyrGluThrGluThrLeuGlyValLeuValLeuValLeuValLeuVal 281
DB 1078 ATGCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
QY 282 PheGluAspHisAspTyrThrLeuGluThrAspSerThrAspSerPheThrGluArg 301
DB 1138 ATGCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
QY 302 ThrProArgLysSerAsnLeuAspGluValValAsnValThrProHisThrProVal 321
DB 1198 AGCAAGCTCAACAAAAGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
QY 322 ArgThrValMetAsnThrThrGluGluLeuMetMetThrLeuAsnSerAlaSerAspGlu 341
DB 1258 AGCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317
QY 342 ProSerGluAsnLeuThrSerTyrPheAsnAspCysThrValAsn 356
DB 1318 ATGCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362

RESULT 2
RM661083 1051 bp mRNA linear EST 20-FEB-2002
DEFINITION ACN5081 5556133 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550405
5', mRNA sequence.
VERSION RM661083
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1051)
AUTHORS NIH MGC Project (2000) GenBank.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

```

```

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: eap@ncbi.nlm.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequenced by: The I.M.A.G.E. Consortium (LLNL)
Cloned Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1264 row: m column: 18
High quality sequence stop: 737.
Location/Qualifiers
1. 1051
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5550405"
/clone_lib="NIH_MGC_67"
/tissue_type="Tetradblastoma"
/lab_host="DH10B (phage resistant)"
/notes="Order: eye; Vector: pMW vector; Stock: 10; Note:
Site 2: Salt; Cloned unidirectionally. Primers: clone dir.
Average insert size 1.75 Kb. Library constructed by Life
Technologies."
BASE COUNT 330 a 221 c 197 g 299 t 4 others
ORIGIN
Alignment Scores:
Pred. No.: 2,046,156 Length: 1051
Score: 1551.00 Matches: 512
Percent Similarity: 93.26% Gaps: 6
Best Local Similarity: 91.50% Mismatches: 18
Query Match: 34.44% Indels: 5
EB: 13 Gaps: 1
US-09-026-459A-31 (1-874) x BM561083 (1-1051)
QY 379 AlaValGlyGluGlyCysValCysLeuThrArgThrCysLeuValArgLeu 398
DB 4 GCTGTGGGACAGGTTGTGTGAAATTTGGATCAGAGGATACAAATTTGGATTCGCTG 64
QY 399 TyrTyrArgValMetGluSerMetLeuLysSerGluGluArgLeuSerThrGluAsn 418
DB 64 TATTACCCAGTAATCGAATCCATGCTTAAATCAACAACAACAACATTAATCATTA 124
QY 419 PheSerLysLeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaLysAlaLeuGlu 438
DB 124 TTATGCAAACTTCGAATGAGAACATTTTCAATGATCTTATTTGGGCTGCTGAG 184
QY 439 ValValMetAlaThrTyrSerArgSerThrSerGluAsnLeuAspSerGlyThrAspLeu 458
DB 184 GTTGTATGGGCAATATAGAGAGAGTAACTCAAACTTGAATGAAATGAAATGAA 244
QY 459 SerPheProTyrPheLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIle 478
DB 244 TCTTCTCCATGCTATGCAATGCTTAAATTTAAAGGCTTCAATTTTCAAACTG 404
QY 479 GluSerPheIleLysAlaGluGlyAsnLeuThrArgGluMetIleLysHisLeuGluArg 498
DB 304 CAATATTTTATCAAGGCAAGAGAGAGTAACTTCAACAACAACAACAATTTTACAG 464
QY 499 CysGluHisArgTyrMetGluSerLeuAlaThrLeuSerAspSerProLeuPheAspLeu 518
DB 364 TGTGAAATCGAATATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
QY 519 IleLysGluSerLysAspArgGluGlyProThrAspHisLeuGluSerAlaLysProLeu 538
DB 424 ATTAACAACATCAAAAGCAAGAGAGAGAGTAACTTCAACAACAACAACAATTTT 484
QY 539 AsnLeuProLeuGluHisAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSer 558
DB 484 AATCTTCTCTCTCAAGAAATTAATCAAGAGAGAGAGTAACTTCAACAACAACAACA 544
QY 559 ProGlyTyrGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGluAla 578

```

Contact: Robert Strausberg, Ph.D.  
Email: eap@ncbi.nlm.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Sequenced by: The I.M.A.G.E. Consortium (LLNL)  
Cloned Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1264 row: m column: 18  
High quality sequence stop: 737.  
Location/Qualifiers  
1. 1051  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5550405"  
/clone\_lib="NIH\_MGC\_67"  
/tissue\_type="Tetradblastoma"  
/lab\_host="DH10B (phage resistant)"  
/notes="Order: eye; Vector: pMW vector; Stock: 10; Note:  
Site 2: Salt; Cloned unidirectionally. Primers: clone dir.  
Average insert size 1.75 Kb. Library constructed by Life  
Technologies."



Accession	PM475603	1001 bp	mRNA	linear	EST 05-FEB-2002
EMBL/GenBank	AC027041	648000	NIH_MOL_92 Homo sapiens cDNA clone IMAGE:5575683		
			6'	mRNA Sequence	

ORGANISM Homo Sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE  
1 (bases 1 to 1061)  
AUTHORS NIH MGC <http://mgc.nci.nih.gov/>

No si sequence available.  
This clone (DKFZp31300816) is available at the RZPD in Berlin.  
Please contact the RZPD: Rossourcenzentrum, Heubnerweg 6, 14059  
Berlin Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

FEATURES
    Source
    1..727
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="DKFZp31300816"
    /clone_lib="313 (synonym: hloc2)"
    /dev_stage="adult"
    /lab_host="F018"
    /note="Vector: pTriplex2; Site: 1..511A; Site 2..511B; cDNA collection"
BASE COUNT      248 a 161 c 130 g 198 t
GC: 31%

Alignment Scores:
Pred. No.:      1,376-124      Length:      727
Score:          1254.00        Matches:      241
Percent Similarity: 99.59%    Conservative: 0
Best Local Similarity: 99.59% Mismatches:      1
Query Match:     27.84%      Indels:       0
DB:              Gaps:       0

US 09-026-459a 31 (1-874) x AL597411 (1-727)

QY 589 CysSerThrSerGlnSerLeuPheTyrLysValTyrAlaGlyAlaTyrGluAlaGlu 608
|||||
DB 2 AAGATGATCTCTTCACTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 61
QY 609 AsnThrLeuGlyAlaGluLeuSerGlnHisPheLeuGluHisLeuLeuPheLeu 628
|||||
DB 62 AATAACCTTCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY 629 LeuPheGluHisThrLeuGluAsnGlyTyrGluLeuMetAlaAspAlaGluHisLeuAspGlu 648
|||||
DB 122 ATTTTCACACACCTTCACACACCTTCACACACCTTCACACACCTTCACACACCTTCACAC 181
QY 649 IleMetMetCysSerMetTyrGlyTyrGlyTyrGlyTyrGlyTyrGlyTyrGlyTyr 668
|||||
DB 182 ATTAATGATGTTTCATATATATATATATATATATATATATATATATATATATATATATAT 241
QY 669 IleIleValThrAlaTyrLysAspLeuProHisAlaValGluLeuThrPheTyrSerVal 688
|||||
DB 242 ATCAATGAAACAGATACAGACAGATACAGACAGATACAGACAGATACAGACAGATACAG 301
QY 689 LeuIleLysGluGluGluTyrAspSerIleLeuValPheTyrAsnSerValPheMetGlu 708
|||||
DB 302 TTGATCAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
QY 709 ArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArgProTyrGlnLeuSerProIle 728
|||||
DB 362 AGAATGAAACAAATATTTTGAATATATATATATATATATATATATATATATATATATAT 421
QY 729 PheHisIlePheArgSerPheTyrLysPheSerSerProLeuArgIleProGlyGly 748
|||||
DB 422 CTCACATATCCCAACACCCCTACACATATCCCTAGTATATATATATATATATATATATAT 481
QY 749 AsnIleTyrIleSerProLeuLysSerProTyrTyrIleSerGlyGlyLeuProThrPro 768
|||||
DB 482 AAATATATATATCCACCTCAAGAGTTCATATATATATATATATATATATATATATATATAT 541
QY 769 ThrLysMetThrProArgSerArgIleIleValSerIleGlySerPheCysThrSer 788
|||||
DB 542 AAAAAATGATCCACATATATATATATATATATATATATATATATATATATATATATATAT 601
QY 789 GluLysPheGlnTyrIleAsnGluMetValCysAsnSerAspArgValLeuLysArgSer 808
|||||
DB 602 CAGAAATATCCCAAAATATATATATATATATATATATATATATATATATATATATATAT 861
QY 809 AlaThrIleLysAsnThrProLysPheLeuTyrLysLeuArgPheAspLeuThrGlySer 828
|||||

```

```

DB 662 GGTGAAGAGAGCAACCCCTTGAACACACTGAAAAAACTACGCTTTGATATTGAAGATCA 721
QY 829 AspGlu 830
|||||
DB 722 CATCAA 727

RESULT 6
BQ222227
LOCUS      BQ222227
DEFINITION AGENCOMPT_756045 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6341563
5', mRNA sequence.
ACCESSION BQ222227
VERSION    BQ222227.1 GI:20403627
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 880)
            NIA M.D. brief comm. 2003 Feb 4/03.
JOURNAL    National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: coapbs-re@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequenced by: Amersham Bioscience Corporation
            Clone distribution: MGC clone distribution. Information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L2AM13279 row, 1 column: 20
            High quality sequence stop: 606.
FEATURES   Location,Qualifiers
            1..880
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:6041563"
            /clone_lib="NIH_MGC_92"
            /tissue_type="embryonal carcinoma, cell line"
            /lab_host="p410A (phage-resistant)"
            /note="Organ: testis, Vector: pCMV-Sport6; Site: 1; Note:
            Size: 880; Cloned unidirectionally; oligo-d primed.
            Average insert size 2.5 kb. Library enriched for
            full-length clones and constructed by life technologies.
            Note: this is a NIH_MGC library."
BASE COUNT 321 a 148 c 152 g 259 t
ORIGIN
Alignment Scores:
Pred. No.:      1,420-122      Length:      880
Score:          1296.50        Matches:      255
Percent Similarity: 92.88%    Conservative: 6
Best Local Similarity: 92.75% Mismatches:     13
Query Match:     27.45%      Indels:       7
DB:              Gaps:       3
US-09-026-459a 31 (1-874) x BQ222227 (1-880)

QY 103 ValLeuPheAlaLeuPheSerLysLeuGluArgThrCysGluLeuIleTyrLeuThrGln 122
|||||
DB 1 GTATTGTTTCATCTTCAAGAAATTGCAAAAGCAATGTAAGCTTATATTATTCACAA 60
QY 123 ProSerSerIleSerIleThrGluIleAsnSerAlaIleValIleLysValSerIlePhe 142
|||||
DB 61 CCAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 120
QY 143 ThrPheLeuLeuAlaLysGlyGluValLeuGluMetGluAspLeuValIleSerPhe 162
|||||
DB 121 ACATTTTATACGTAAGGCAAGTATTACAAA GCAAGATCACTGGTGATTTTCA 180
QY 153 GluLeuMetLeuGlyValLeuAspTyrPheIleLysLeuSerPheProMetLeuLys 182
|||||

```











```

KEYWORDS: EST,
SOURCE: human.
ORGANISM:
REFERENCE:
AUTHORS:
TITLE:
JOURNAL:
MEDLINE:
COMMENT:
SEQUENCE:
LOCUS:
DEFINITION:
ACCESSION:
VERSION:
KEYWORDS:
SOURCE:
ORGANISM:
REFERENCE:
AUTHORS:
TITLE:
JOURNAL:
COMMENT:
FEATURES:
SOURCE:

```

432 LeuLeuAlaGlyValAlaGlyValAlaThrTyrSerAlaGlyThrSerGluAsn 431  
 437 TTAATGGGATGGCTCTTCAAGCTTCTGAAAGGAGCAATAGAAAGAAATACAA 436  
 452 LeuAspSerGlyThrAspLeuSerPheProIleIleGluAsnValLeuAsnLeuGlyAla 471  
 477 CTGATTCGGAAAGATTAATGCTTTTTCATGGATTTCGAAAGGCTTAATTAAGAAAG 438  
 472 PheAspPheCysValLeuGluSerPheCysValLeuGluSerPheCysValLeuGluSer 491  
 437 TTTGATTTTAAATGATGCAAGATTTTAAATGATGCAAGATTTTAAATGATGCAAG 436  
 492 MetLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 511  
 257 ATGATAAACAATTAAGAAAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAG 198  
 512 AspSerProLeuLeuAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 541  
 197 GATTCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 148  
 532 LeuGluSerAlaGlyProLeuAsnLeuProLeuGluSerAlaGlyProLeuAsnLeu 551  
 137 CTGAAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 78  
 551 TTTTLeuSerProValAlaGlySerProLeuGlySerThrThrAlaVal 468  
 77 GTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 26  
 624 bp. mRNA linear EST 13 MAR 2002  
 IMAGE:5638049 57 similar to db:MI5400 RETINOBLASTOMA ASSOCIATED  
 PROTEIN (HUMAN); db:M26391 Mouse retinoblastoma susceptibility  
 protein (MOUSE); mRNA sequence.

ACCESSION: AW584181.1 GI:7260125  
 VERSION: AW584181.1  
 KEYWORDS: EST.  
 SOURCE: human.  
 ORGANISM: Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 624)  
 Melton, D., Meadows, A., Clifton, S., Hillier, L., Marra, M., Pape, B.,  
 Wyte, F., Maitin, J., Blistain, A., Schmitt, A., Theisling, R., Rutter,  
 R., Bonko, L., Bennett, J., Cardenas, M., Gibbons, M., McAnney, K., Cule,  
 R., Tsagaris, W., Williams, L., Jackson, Y., and Bowers, Y.  
 WashU-Harvard Pancreas EST Project  
 Unpublished (2000)  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@bioapp.harvard.edu  
 Libraries were constructed by Dr. Douglas Melton  
 DNA sequencing by: Washington University Genome Sequencing Center  
 For information on obtaining a clone please contact: Juliana Brown  
 (brownj@wustl.edu)  
 This sequence was available from the IMAGE Consortium. For clone  
 orders contact: info@image.llnl.gov  
 Seq primer: -40bp from Gibco  
 High quality sequence stop: 430.  
 Location/Qualifiers  
 1..624  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5638049"  
 /clone\_lib="human Pancreatic Islets"  
 /issue\_type="Islets of Langerhans"  
 /dev\_stage="Adult"





Tue Jan 21 10:02:53 2003

us-09-026-459a-31.rst

Page 14

Search completed: January 19, 2003, 05:05:33  
Elap time : 1945.65 secs

---

Genome version 5.1.4  
Copyright (c) 1994 - 2003 Computer Ltd.

us-09-026-459a-32

Search time: 291.86 seconds  
(without alignments)  
18469,506 Million cell updates/sec

us-09-026-459a-32

Sequence: 1 GATTAATGAGAGAAAC11102.....AAATGAGATTAATGAACT 3323

Scoring table: IDENTITY: 100  
Gap: 10.0, Gapext: 1.0

Search: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42409142

Minimum hit seq length: 0

Maximum hit seq length: 200000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

Listed first 45 summaries

Database: EST

- 1: cm\_estb1\*
- 2: cm\_estb1\*
- 3: cm\_estb1\*
- 4: cm\_estb1\*
- 5: cm\_estb1\*
- 6: cm\_estb1\*
- 7: cm\_estb1\*
- 8: cm\_estb1\*
- 9: cm\_estb1\*
- 10: cm\_estb1\*
- 11: cm\_estb1\*
- 12: cm\_estb1\*
- 13: cm\_estb1\*
- 14: cm\_estb1\*
- 15: cm\_estb1\*
- 16: cm\_estb1\*
- 17: cm\_estb1\*
- 18: cm\_estb1\*
- 19: cm\_estb1\*
- 20: cm\_estb1\*
- 21: cm\_estb1\*
- 22: cm\_estb1\*
- 23: cm\_estb1\*
- 24: cm\_estb1\*
- 25: cm\_estb1\*
- 26: cm\_estb1\*
- 27: cm\_estb1\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.2	27.4	1024	BM450031	BM450031 AGENT0001
2	909.4	27.4	1024	BM450031	BM450031 AGENT0001
3	896.6	25.8	1091	BQ439221	BQ439221 AGENT0001
4	844.8	25.4	881	BQ222275	BQ222275 AGENT0001
5	824.6	24.8	1364	AK011246	AK011246 Mus muscu
6	786	24.7	1001	BM475603	BM475603 AGENT0001

7	752.4	22.6	1122	13	BM450031	BM450031 AGENT0001
8	725.4	21.8	727	9	AL597911	AL597911 DKE2P31.40
9	720.4	21.7	882	14	BQ222275	BQ222275 AGENT0001
10	699.6	21.1	823	12	BQ610661	BQ610661 602611148
11	698.2	21.0	792	9	AL564757	AL564757 AL564757
12	692.2	20.8	747	12	BQ504356	BQ504356 602611142
13	691.8	20.8	972	13	BM466194	BM466194 AGENT0001
14	678	20.4	835	13	BM453724	BM453724 AGENT0001
15	660	19.9	672	13	B1868359	B1868359 603492480
16	643.6	19.4	797	10	BQ549278	BQ549278 601063567
17	642.2	19.4	681	12	BQ616219	BQ616219 602444717
18	617	18.6	618	13	BM264009	BM264009 30341081X
19	614	18.5	694	10	BQ082836	BQ082836 RC2 H1604
20	605.2	18.2	623	10	AW583181	AW583181 30111091Y
21	605	18.2	605	9	AL599633	AL599633 DKE2P31.40
22	596.4	17.9	922	12	BQ253543	BQ253543 602464278
23	580	17.5	580	9	AL599105	AL599105 DKE2P31.40
24	575.2	17.3	914	9	AA764411	AA764411 VW54404.1
25	566.4	17.0	588	10	AW368234	AW368234 CM3 H161R
26	558	16.8	871	9	AA764485	AA764485 VW54404.1
27	535.2	16.0	598	10	AW582967	AW582967 1006080Y
28	530	15.9	539	9	AL692790	AL692790 DKE2P31.40
29	528	15.9	522	9	AL135943	AL135943 DKE2P31.40
30	516.8	15.6	597	12	BQ541246	BQ541246 860 H161C
31	513.2	15.4	594	10	AV715533	AV715533 AV715533
32	512.2	15.4	734	12	B1131957	B1131957 60179043X
33	505.2	15.2	736	13	B1131568	B1131568 602611148
34	493.8	14.9	525	9	AA268255	AA268255 2159004.1
35	486.4	14.6	584	12	BQ149050	BQ149050 0086001Y
36	484.4	14.6	494	10	AW502887	AW502887 01 H1600
37	481.4	14.5	484	10	B1168095	B1168095 QV3 H161
38	480.2	14.5	659	13	B1821116	B1821116 603035591
39	476.6	14.3	2083	13	BM456728	BM456728 AGENT0001
40	474.4	14.3	734	13	B1526982	B1526982 602927177
41	459.2	13.8	497	12	BQ44784	BQ44784 RC3 H1604
42	457	13.8	457	9	AL094215	AL094215 qB04402.X
43	457	13.8	732	9	AL598766	AL598766 DKE2P31.40
44	454.2	13.7	869	12	BQ053749	BQ053749 601678141
45	447.2	13.5	832	9	AL646048	AL646048 VW54404.1X

ALIGNMENTS

RESULT 1  
BM450031  
LOCUS AGENT0001\_6393544 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5528037  
DEFINITION 5' mRNA sequence.  
ACCESSION BM450031  
VERSION BM450031.1 GI:18459071  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (Bases 1 to 1023)  
AUTHORS NIH-MGC <http://mrc.nhlbi.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [rausberg@mail.nih.gov](mailto:rausberg@mail.nih.gov)  
Tissue Procurement: ATCC/DCT10/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM1204 Row: 1 Column: 22  
High quality sequence stop: 643  
Location/Qualifiers 1..1023



```

121  AATTIAGCAAAATTCTGATGAAAGAAATTTTATATGCTTTATATGAGGTGGGTCTT 180
QY  1249  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1308
DB  1249  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1308
DB  181  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 240
QY  1309  TTGCTTTCTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368
DB  1309  TTGCTTTCTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368
DB  241  TGGCTTTCTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY  1369  ATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTT 1428
DB  1369  ATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTT 1428
DB  301  ATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTTATCAAAAGCTTT 360
QY  1429  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1488
DB  1429  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1488
DB  421  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 480
QY  1549  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1608
DB  1549  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1608
DB  481  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 540
QY  1609  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1668
DB  1609  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1668
DB  541  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 600
QY  1669  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1728
DB  1669  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1728
DB  601  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 660
QY  1729  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1788
DB  1729  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1788
DB  661  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 720
QY  1789  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1848
DB  1789  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1848
DB  721  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 780
QY  1849  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1908
DB  1849  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1908
DB  781  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 840
QY  1909  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1967
DB  1909  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 1967
DB  841  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 900
QY  1968  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 2019
DB  1968  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 2019
DB  901  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 963
RESULT 3
QY  1968  CAGGCTTCTTAAAGGCAATATATAGAGAGACATATGACAAATCTGATATGAGAAACAGAT 2019
DEFINITION  AGN_0001 776,562 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015480
5% mRNA sequence.
VERSION  BQ439221
KEYWORDS  EST
SOURCE  human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS  NIH MGC http://mgi.nih.gov/
TITLE  National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
Email: rstraus@mail.nih.gov

```

Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)  
cDNA Sequencing by: Agencourt Bioscience Corporation (LLNL)  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM14211 row: k column: 01  
High quality sequence step: 742.  
location/Qualities  
1: 1081  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="6015480"  
/clone\_lib="NIH\_MGC\_92"  
/class="type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: pCMV Sp6b16; Site: 1; Note:  
Site\_2: Salt; cloned unidirectionally; cloned at primed.  
Average insert size: 2.5 kb. Library enriched for  
full length clones and constructed by life technologies.  
Note: this is a NIH\_MGC library."

BASE COUNT 337 a 190 c 183 q 471 t  
ORIGIN  
Query Match 25.8% Score 856; DB 14; Length 1081;  
Best local Similarity 99.1% Pred. No. 8,7e-155;  
Matches 892; Conservative 0; Mismatches 5; Gaps 42

```

QY  2425  GAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2484
DB  2425  GAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2484
DB  1  GAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY  2485  ATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2544
DB  2485  ATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2544
DB  61  ATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY  2545  AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2604
DB  2545  AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2604
DB  121  AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY  2605  CATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2664
DB  2605  CATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2664
DB  181  CATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2724
QY  2665  ATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2724
DB  2665  ATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2724
DB  241  ATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY  2725  AAGTCACCTTGAATTTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2784
DB  2725  AAGTCACCTTGAATTTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2784
DB  301  AAGTCACCTTGAATTTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY  2785  TGGCATTTTAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2844
DB  2785  TGGCATTTTAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2844
DB  361  TGGCATTTTAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY  2845  TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2904
DB  2845  TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2904
DB  421  TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY  2905  TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2964
DB  2905  TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2964
DB  481  TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY  2965  TTTTCTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3024
DB  2965  TTTTCTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3024
DB  541  TTTTCTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY  3025  TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3084
DB  3025  TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3084
DB  601  TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

```



















[illegible][illegible]











QY 82 PheAlaLeuPheSerLysLeuGluAlaArgThrCysGluLeuLleIleThrLeuThrGlnProSer 101  
 DB 607 TGTGACACTCAACGAAATTAACAGGAGGTGTGAATTAATATATATTTGACACAGCCAGC 666  
 QY 102 SerSerLleSerPheGluLleAsnSerAlaLeuValLeuLysValSerTrpIleThrPhe 123  
 DB 667 AGTGAGTATACGATGAGAAATTAATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 726  
 QY 122 LeuLeuAlaLysGlyGluValLeuGluMetGluAspSerLeuValLleSerPheGluLeu 141  
 DB 727 TTAATAGTAAAGACACAAATTAACAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 786  
 QY 142 MetLeuLysValLeuAspTrpPheLleLysLeuSerProMetLeuLysGluPro 161  
 DB 787 ATGT 846  
 QY 162 TyrLysThrAlaValLleProLleAsnGlySerProArgThrProAlaArgGlyGluAsn 181  
 DB 847 TACAAATAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906  
 QY 182 ArgSerAlaArgLleAlaLysLeuGluAsnAspThrArgLleLleGluValLeuLys 201  
 DB 907 AGAGAGGTCGCAATGACAAATTAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966  
 QY 202 LysLeuLysSerLysAsnLleAspGluValLysAsnValTrpPheLysAsnProLlePro 221  
 DB 967 AAGAGAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026  
 QY 222 PheMetAsnSerLeuGlyLeuValLleSerAsnGlyLeuProGluValGluAsnLeuSer 241  
 DB 1027 TTTAATAATGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086  
 QY 242 LysArgTyrGluGluLleTyrLeuLysAsnLysAspLeuAspAlaArgLeuPheLeuAsp 261  
 DB 1087 AAAAGCTATGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1146  
 QY 262 HisAspLysThrPheGluThrAspSerLleAspSerPheGluThrGluArgThrProArg 281  
 DB 1147 CATGATAAATATGACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1206  
 QY 282 LysSerAsnLeuAspGluGluValAsnValLleProProHisIleThrProValArgThrVal 301  
 DB 1207 AAAAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1266  
 QY 302 MetAsnThrLleGluGluLeuMetMetLleLeuAsnSerAlaSerAspGluProSerGlu 321  
 DB 1267 ATGAAATATATCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1326  
 QY 322 AsnLeuLleSerTyrPheAsnAspCysThrValAsn 341  
 DB 1327 AATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 1382  
 REFSeq 2  
 RM561083  
 LOCUS 1051 bp mRNA linear EST 20 FEB 2002  
 DEFINITION AGEM000001.656614 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:5550905  
 5' mRNA sequence.  
 ACCESSION RM561083  
 VERSION RM561083.1 GI:19809043  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1051)  
 AUTHORS NIH MGC <http://mgi.scripps.edu/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [rapb@scripps.edu](mailto:rapb@scripps.edu)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Amersham Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LIML at:  
<http://image.llnl.gov>  
 Plate: LAM12263 row: m column: 18  
 High quality sequence stop: 737.  
 Location/Qualifiers  
 1..1051  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5550905"  
 /clone\_lib="NIH\_MGC\_67"  
 /issue\_type="retroblastoma"  
 /lab\_host="B10.108 (phage-resistant)"  
 /note="Organ: eye; Vector: pCMV SPORT6; Site: 1; Note:  
 Site\_2: Salt; Cloned unidirectionally. Primer: oligo dT.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies."

BASE COUNT 330 a 221 c 197 g 295 t 4 others  
 ORIGIN

Alignment Scores:  
 Prod. NO.: 155e-154 Length: 1051  
 Score: 1551.00 Matches: 412  
 Percent Similarity: 93.25% Ins. Penalty: 6  
 Best Local Similarity: 91.50% Mismatches: 18  
 Query Match: 45.48% Indels: 5  
 DB: 13 Gaps: 1

US-09-026-459A-33 (1-851) x HM561083 (1-1051)

QY 356 AlaValGlyGlnGlyCysValGluLleGlySerGlnArgTyrLysLeuGlyValArgLeu 475  
 DB 4 GTGTGGACACAGGCTG 63  
 QY 376 TyrTyrArgValMetGluSerMetLeuLysSerGluGluGluArgLeuSerLleGluAsn 395  
 DB 64 TATTATCGAGTAATGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 123  
 QY 396 PheSerLysLeuLeuAsnAspAsnLlePheHisMetSerLeuLeuAlaLysAlaLeuGlu 415  
 DB 124 TTTACCAAACTTCGAAACACACAACTTTTCAATGCTTTATTCGAGGAGGAGGAGGAG 184  
 QY 416 ValValMetAlaThrTyrSerArgSerThrSerGlnAsnLleAspSerGlyThrAspLeu 445  
 DB 184 GTTGTAAATGGGCAATATAGAGAGATCAATGAGAAATGTGATTTGAGAAAGATTG 243  
 QY 436 SerPheProTrpLleLeuAsnValLleuAsnLeuLysAlaPheAspPheCysValLle 455  
 DB 244 TCTTTCCCATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 404  
 QY 456 GluSerPheLleLysAlaGlnLysAsnLeuThrArgGluMetLleLysHisLeuGluArg 475  
 DB 304 GAAATCTTTATCAAGACCAACGCAACCTTCGAAATGATGATGATGATGATGATGATGAT 404  
 QY 476 CysGluHisArgLleMetGluSerLeuAlaLleTrpLeuSerAspSerProLeuPheAspLeu 495  
 DB 464 TGTGAACATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424  
 QY 496 LleLysGlnSerLysAspArgGluGlyProThrAspHisLleGluSerAlaLysProLeu 515  
 DB 424 ATTAACAAATCAAGAGACGAGAGAGACAACTGATATATGATGATGATGATGATGATGAT 484  
 QY 516 AsnLeuProLeuGlnAsnAsnLleThrAlaAlaAspMetTyrLeuSerProValArgSer 545  
 DB 484 AATCTCTCCCTCCCAATTAATCAGACATGACAGCAGATATGATGATGATGATGATGATG 544  
 QY 536 ProLysLysLysGlySerThrThrArgValAsnSerThrAlaAsnArgLleGlnAla 555  
 DB 544 CCAAG 603  
 QY 556 ThrSerAlaPheGlnThrGlnLysProLeuLysSerThrSerLeuSerLlePheLysLys 575

```

DB 604 AACCTAGCTTCCACACACACAGACATGCAAAATACCTCTCTTACACGTGTTTAAAA 663
QY 676 LysValTyrArgLeuAlaTyrLeuArgLeuAsnThrLeuCysGluArgLeuSerGlu 595
DB 604 AAATGTAATGCTAGCTTATGCTGGGTAAATACATTTCTGAAAGGCTCTGCTCTAG 723
QY 676 HisProGluGluGlnHisLeuTyrPheThrCysPheLeuHisThrLeuLeuAsnGlyTyr 615
DB 724 CACCGAGATTAGACATATCATCTGAGACCTTTTCCAGACACACGUGAGAAAGATAT 783
QY 676 GluGlnGlnTyrAspArgHisLeuAspGlnLeuMetMetCysSerMetTyrGlyTyrCys 635
DB 784 GAACATATACACACACACACACACANTTTCAGACCAAAATATGATGATGCTCCATAGG 843
QY 676 LysValTyrAspLeuAspLeuLysPheCysGluLeuValTyrLeuTyrLysAspLeu 655
DB 844 AAAGTGAAGATATAGACATTTAAATATTCGAAATGCTTGTAAACAGCATATACAGATCT 903
QY 676 HisAlaValGlnGluThrProGlyArgValLeuTyrLeuTyrGlnGluGlnTyrAspSer 674
DB 904 TCATCTGTTNTGAGACACATTTAAAGCTGTGTTGATCCAAAGAGAGAGATATGATG 963
QY 676 IleIleVal-----PheGlyTrpSerSerValIlePheMetGluArgLeuGlyMetAsp 691
DB 904 TATATACGATATGATATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

RESULT 4
LOCUS BQ220275 881 bp mRNA linear EST 02-MAY-2002
DEFINITION AF090007.7572421 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575683
5', mRNA sequence.
ACCESSION BQ220275
VERSION BQ220275.1 GI:20401675
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 881)
AUTHORS NIH-MGC http://mgi.nhl.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: eadp@small.nhl.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrived by: the I.M.G.E. Consortium (LIN)
DNA Sequencing by: Advanced Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LIN at:
http://linear.lin.gov
Plate: LLAM13287 row: e column: 03
High quality sequence stop: 690.
FEATURES
location/Qualifiers
1..881
/organism="Homo sapiens"
/clone="IMAGE:5575683"
/clone="IMAGE:604522"
/tissue_type="embryonal (phage-resistant)"
/label="organ: testis; Vector: pCMV-Sport6; Size: 1.1 kb; Note:
Site 3: Salt: cloned and directly: clone of primed
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies
Note: this is a NIH_MGC library."
BASE 313 a 149 c 155 g 264 t
GC 13.1%

```

Alignment Scores:

Prod. No.: 1.17e-143

Length: 881

Matches: 1448/00

```

Percent Similarity: 98.40% Conservative: 1
Best Local Similarity: 97.96% Mismatches: 3
Query Match: 33.12% Indels: 2
DB: 14 Gaps: 0
US-09-026-459a-33 (1-881) x BQ220275 (1-881)
QY 131 MctCHeAspSerLeuValTyrSerPheGlnMetLeuCysValLeuAspTyrPhe 150
DB 3 ATGAAGAATGATGCGGATGATTTATTTCAATTAATGCTATGCTGCTGCTGCTGCTGCT 62
QY 151 LysLeuSerThrProMetLeuLeuGlySerGluTyrGlyTyrThrAlaValTyrGlnLeu 170
DB 63 AAATGCTACAGTGGCACTTGGTCAACCAACCAATATAAAACACCTCTATACCCAATAA 122
QY 171 GlySerThrArgThrPheArgLeuArgGlyGluAsnArgSerAlaArgLeuAlaLysGlnLeu 190
DB 123 GCTTCACCTTGAAACACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 182
QY 191 GluAsnAspThrArgTyrLeuGluValLeuCysTyrSerLeuHisGlnCysAspHisCasp 210
DB 183 GAAATGATATACGATATATTTCAAGTTCTCTATTAATTAATTAATTAATTAATTAATTA 242
QY 211 ValLysAsnValTyrPheLysAspPheLeuTyrMetTyrSerLeuGlyValThr 230
DB 243 CTGAAAATGCTTATTTCAAAAATTTTATACCTTTATCAATTTCTGCTGCTGCTGCTGCT 302
QY 231 SerAspGlyLeuProGluValGluAsnLeuSerLysArgTyrGluGlnLeuTyrLeuLys 250
DB 303 CTAAAGCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 362
QY 251 AsnLysAspLeuAspAlaArgGlnCysPheLeuAspHisAspLysThrLeuGlnThrAspSer 270
DB 363 AATAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 422
QY 271 IleAspSerPheGluThrGlnArgThrProArgLysSerAsnLeuAspGluLeuValAsn 290
DB 423 ATACACACTTCAACACACACACACACACACACACACACACACACACACACACACACACAC 482
QY 291 ValIleProPheThrPheValArgThrValMetLeuAsnThrLeuGlnLeuMetMet 310
DB 483 GTAAATCTCCACACACACACACACACACACACACACACACACACACACACACACACAC 342
QY 311 IleLeuAsnSerAlaSerAspGlnProSerGlnAsnLeuSerLysThrPheAsnAsnGly 330
DB 543 ATTTTAAATTCACCAATATGATCAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 402
QY 331 ThrValAspThrGlnSerLysLeuLysArgValGlyAspPheGlyTyrPheLys 350
DB 603 ATATGATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 462
QY 451 HisLysPheAlaLysAlaValGlyGlnTyrGlyValLeuTyrLeuGlySerGlnArgTyrLys 470
DB 663 GAGAAATTTCTTAAAGCTGCTGAGACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
QY 371 LeuGlyValArgLeuTyrTyrArgValMetGluSerMetLeuLysSerGluLeuGluArg 390
DB 723 GTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
QY 391 LeuSerLeuGlnAsnPheSerLysLeuAspHisAspHisPheHisMetLeuLeu 410
DB 783 TTAATCAATCAAAATTTTACCAAACTTTCGAAATTCGAAATTCGAAATTCGAAATTCGAA 842
QY 410 GluAlaCysAlaLeuGluValMetLeuTyrTyrSer 422
DB 843 TGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
RESULT 4
BM475603
LOCUS BM475603 1001 bp mRNA linear EST 05-FEB-2002
DEFINITION AF090007.7572421 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575683
5', mRNA sequence.
ACCESSION BM475603

```













```

CY 243 ArgTyrGluGluIleTyrLeuLysAsnLysAspAlaValLeuPheLeuAspVal 262
      |||||
CY 244 GAAAGCAAGCAAAATTTTCTTAAAAATACAACTACATCAAGAGATATTTTGGTTCAT 492
      |||||
CY 245 AspLysThrLeuGlnThrAspSerLeuAspSerPheGluThrGlnArgThrProArgLys 282
      |||||
DB 491 GATATAAAATGTTTASATATATTGATGAGATGTTTAAATGATGATGATGATGATGAT 442
      |||||
CY 246 SerAsnLeuAspGluGluValAsnValLeuProGlnHisThrProValArgThrValMet 302
      |||||
DB 493 AGTAAACCTGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372
      |||||
CY 403 AsnThrIleGlnLeuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsn 322
      |||||
DB 471 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
      |||||
CY 423 LeuLysSerThrPheAsnAsnGlnThrValAsnProLysCysSerIleLeuLysArgVal 342
      |||||
DB 494 YTGTTTCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252
      |||||
CY 443 LysAspIleGlyTyrIlePheLysGluLysPheAlaValGluGlnGlyCysVal 362
      |||||
DB 251 AAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 192
      |||||
CY 463 GluIleGlySerGlnArgTyrLysLeuGlyValArgLeuTyrTyrArgValMetGlu 382
      |||||
DB 495 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 132
      |||||
CY 482 GlnMetLeuLysSerGluGlnGluGluGluGluGluGluGluGluGluGluGluGlu 401
      |||||
DB 496 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72
      |||||
CY 402 AspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrTyr 421
      |||||
DB 71 ACAAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12
      |||||
CY 422 SerArgSer 424
      |||||
DB 11 GATGATGAT 3
      |||||
FEATURES
     AL599633
     DEFINITION DKF7314P092.1 313 (synonym: hlec2) Homo sapiens cDNA clone
     VERSION DKF7314P092.1 mRNA sequence.
     KEYWORDS AL599633.1 GI:15162921
     SOURCE FST
     ORGANISM human.
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo
     (bases 1 to 605)
     REFERENCE
     AUTHORS Koehler,K., Beyer,A., Mewes,W., Weill,B. and Wiemann,S.
     TITLE EST (Koehler,K., Beyer,A., Mewes,W., Weill,B. and Wiemann,S.)
     JOURNAL Unpublished (1999)
     COMMENT Contact: Koehler K
     MISC
     Am Klopferspitze 18a D 92152 Martinsried, Germany
     This is the 5' sequence of the clone insert.
     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
     Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
     sequenced by HNFZ (Biomedical Research Center at the Charité,
     Berlin/Hamburg) within the cDNA sequencing consortium of the German
     Genome Project.
     No sl sequence available.
     This clone (DKF7314P092) is available at the EFG in Berlin.
     Please contact the RZPD, Bessourceszentrum, Heubnerweg 6, 14059
     Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
     LOCATION/Qualifiers
     Locus: 605 bp
     ORGANISM: Homo sapiens
     /db_xref "taxon:9606"

```

```

/clone="DKF7314P092"
/clone_lib="313 (synonym: hlec2)"
/dev_stage="adult"
/lab_host="PH10H"
/notes="Vector: pTIP10x2 Site_1: Still; Site_2: Still;
cDNA-collection"
BASE COUNT 228 a 92 c 106 q 179 t
ORIGIN
Alignment Scores: 4.84e-100 Length: 605
Pred. No.: 1037.00 Matches: 201
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 23.72% Gaps: 0
DB: 9
US-09-026-459a-33 (1-851) x AL599633 (1-605)
QY 198 GlnValLeuGlyLysGluHisGluCysAsnIleValSerGlnValLysAsnValLysPheLys 217
      |||||
DB 1 GAAGTCTCTCTGTAAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
      |||||
QY 218 AsnPheIleProPheMetAsnSerLeuGlyIleValThrSerAsnGlyLeuProGluVal 237
      |||||
DB 61 AATTTTATATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
      |||||
QY 238 GluAsnLeuSerLysArgTyrGluGlnGlyCysLeuLysAsnLysAspLysAspLysVal 257
      |||||
DB 121 GAAATCTCTCTGTAAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
      |||||
QY 258 LeuPheLeuAspHisAspLysThrLeuGlnThrAspSerIleAspSerPheGluThrGln 277
      |||||
DB 181 TATTTTGTATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
      |||||
QY 278 ArgThrProArgLysSerAsnLeuAspGluGluValAsnValIleProProIleThrPro 297
      |||||
DB 241 GAAACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
      |||||
QY 298 ValArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAsp 317
      |||||
DB 301 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
      |||||
QY 318 GlnProSerGlnAsnLysIleSerIlePheAsnAspCysThrValAsnProLysGluSer 437
      |||||
DB 361 CAACCTTCAGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
      |||||
QY 338 IleLeuLysArgValLysAspIleGlyTyrIlePheLysGluLysPheAlaLysAlaVal 357
      |||||
DB 421 ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
      |||||
QY 358 GlyGlnGlyCysValGlnGlyGlySerGlnValTyrIleLysGluGlyValArgThrVal 377
      |||||
DB 481 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
      |||||
QY 378 ArgValMetGluSerMetLeuLysSerGluGluGluArgLeuSerIleGlnAsnPheSer 497
      |||||
DB 541 CGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
      |||||
QY 398 Lys 398
      |||||
DB 601 AAA 603
      |||||
RESULT 12
BE082846/c
LOCUS BE082846
DEFINITION BE082846 Homo sapiens cDNA, clone 42, 694 bp
ACCESSION BE082846
VERSION BE082846.1 GI:8473151
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

BE082846 694 bp mRNA linear EST 12-JUN-2000  
 BE082846 Homo sapiens cDNA, clone 42, 694 bp  
 BE082846.1 GI:8473151  
 EST.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



















DB 723 CTGGACAGTGGTGTGATTAACAGCAAGAAAGCAATGCTGTAATGCAAGCAAGCAAGCA 782  
 QY 468 LeuSerThrGluAspHisSerLysLeuLeuAspAsnThrPheHisMetSerLeuL 487  
 DB 785 TATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 842  
 QY 487 GAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499  
 DB 845 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880  
 RESULT 4  
 QY 468 1001 bp mRNA Linear EST 05-FEB-2002  
 DEFINITION AGEM0001.640007 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5575683  
 5' mRNA sequence.  
 ACCESSION U0475603  
 VERSION U0475603.1 GI:18524645  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1001)  
 AUTHORS NIH MGC <http://mhc.nhl.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [rcs@mail.nih.gov](mailto:rcs@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA library Preparation: Life Technologies, Inc.  
 cDNA library Arrayed by: the I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Placed in AM12426 row 1 column 04  
 High quality sequence starts 17  
 High quality sequence stops 744  
 Location/Qualifiers  
 source  
 Zorganism="Homo sapiens"  
 Zdb\_xref="taxon:9606"  
 Zclone="IMAGE:5575683"  
 Zclone\_lib="NIH\_MGC\_92"  
 Zissue\_type="embryonal carcinoma, cell line"  
 Zlab="bill63 (phase-resistant)"  
 Znote="Random testis: Vector: pCMV sport6; Site\_1: Not I;  
 Site\_2: SalI; cloned unidirectionally; oligo dI primed;  
 Average insert size 2.5 kb. Library enriched for  
 full length clones and constructed by life technologies.  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 449 a 178 c 185 g 288 t  
 ORIGIN  
 Alignment Statistics  
 Seq. No. 4 140 124 Length: 1001  
 Scores 1439.80 Matches: 281  
 Percent Similarity 99.87% Conservatism: 10  
 Best Local Similarity 99.85% Mismatches: 10  
 Query Match 27.91% Indels: 9  
 Gaps: 3  
 us-09-026-459a-2 (1-929) x U0475603 (1-1001)  
 QY 135 LeuLysGluLeuAspHisSerThrLysValAspAsnAlaMetSerAlaLeuLeuLysLys 154  
 DB 48 GTCGCAAGCAAGCAATATAAAGACGCTGTTATATGCTGCTGCTGCTGCTGCTGCTGCT 497  
 QY 155 TCAATTCAGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 174  
 DB 98 TATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 157  
 QY 174 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 194

DB 158 GACACAAATGACAGTGTGATTAACAGCAAGAAAGCAATGCTGTAATGCAAGCAAGCA 217  
 QY 194 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 214  
 DB 218 TTGGATTCACATTTTATAGCTAAGGCAAGCAATATAAAGACGCTGTTATATGCTGCTGCT 277  
 QY 214 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 244  
 DB 278 TTCAATTCAGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 347  
 QY 234 GTCGCAAGCAAGCAATATAAAGACGCTGTTATATGCTGCTGCTGCTGCTGCTGCTGCT 344  
 DB 338 GTCGCAAGCAAGCAATATAAAGACGCTGTTATATGCTGCTGCTGCTGCTGCTGCTGCT 397  
 QY 254 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 274  
 DB 398 GTCGCAAGCAAGCAATATAAAGACGCTGTTATATGCTGCTGCTGCTGCTGCTGCTGCT 457  
 QY 274 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 294  
 DB 458 TCAATTCAGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 517  
 QY 294 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 314  
 DB 518 AAATTTTATACCTTTTATGCAATTTCTTGGACCTTCTAATCATCTAATGCTATTTTAAAG 577  
 QY 314 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 334  
 DB 578 TCAATTCAGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637  
 QY 334 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 354  
 DB 638 ATTATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 697  
 QY 354 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 374  
 DB 698 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 757  
 QY 374 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 394  
 DB 758 AGTACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 817  
 QY 394 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 413  
 DB 818 TCAATTCAGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877  
 QY 413 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 429  
 DB 878 AAATTTTATACCTTTTATGCAATTTCTTGGACCTTCTAATCATCTAATGCTATTTTAAAG 937  
 QY 430 GATGATTCAGAAATTTTGAATTTGGAATGACACATTTTTCATATGTCCTTAT 445  
 DB 938 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957  
 RESULT 5  
 QY 430 1001 bp mRNA Linear EST 14-APR-2002  
 DEFINITION DKFZP4340816.1 434 (576-676) Homo sapiens cDNA clone  
 DKFZP4340816.1, mRNA sequence.  
 ACCESSION U0475603  
 VERSION U0475603.1 GI:15160502  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 727)  
 AUTHORS Rouska, A., Wellenreuther, R., Mewes, H.W., Well, B., and Wiemann, S.  
 TITLE EST (Rouska, A., Wellenreuther, R., Mewes, H.W., Well, B., and Wiemann, S.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Poustka, A.J.





```

650 AspSerIleLeValPheTyrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeu 769
      |||
608 CAlICIAIAAGIAIACCTAIAACTCGCTCTTCATCCAGACAGACGAAACAAATATTC 567
      |||
670 GlnTyrAlaSerThrArgProProThrLeuSerProIleProHisIleProArgSe 788
      |||
668 CAGTAIGATTGACAGACGCGCGCTAGCGTTGTGCACCAANTAGCTCACAATTTCCTGGA 727
      |||
698 rProTyrLys---PheProSerProLeu---ArgIlePro---GlyGlyAsn----- 803
      |||
628 GCGGCTTAAAGGTTCCCTCACTTACGCAACCTTACGCAACCTTACGCAACCTTAC 787
      |||
604 -IleTyrIleSerProLeuLys 810
      |||
688 TAlAlTTCACACCCCTCGGAA 809

RESULT 8
RM450031 1023 bp mRNA linear EST 05-FEB-2002
DEFINITION A5EN00PT 6303544 NIH MGC 72 Homo sapiens cDNA clone IMAGE:5528037
5' mRNA sequence.
A5EN00PT
VERSIN RM450031.1 GI:18499071
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1023)
NIH-MGC http://mgs61.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
AUTHOR CONTACT: Robert Strausberg, Ph.D.
EMAIL: eugene.femail.nih.gov
HSSOC Procurement: A5CC/UCI/ULP
cDNA Library Preparation: Life Technologies, Inc
DNA Sequencing by Applied Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: L1AM02204 row 3 column 22
High quality sequence stop: 643.
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528037"
/clone_lib="NIH-MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/vec="pCRan SKin Vector pCMV SKN2000 Site 1 Ref.
Site 2 Ref. Cloned unidirectionally. Primary Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 444 a 211 c 197 g 281 t
GC 39.26
Alignment Scores:
Prod. No. 2 76-100 Length 1077
Score 1111.00 Matches 216
Percent Similarity 100.00% Conservations 0
Best local Similarity 100.00% Mismatches 0
Query Match 23.16% Indels 0
Gap 13 Gaps 0
23 26 459a * (1 428) * BM450031 (1-1023)
674 LysValPheAsnIleLeValPheTyrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeu 732
      |||
674 AAAGTGAAGAAATACAGCTTAAATTCACCAATTCATGTCACAGCAACACAGATCTCTCT 61
      |||
674 HisAlaValGlnGlnThrPheLysAsnValLeuIleLysGlnGlnGlnGlnGlnGln 752

```

```

      |||
62 CAlICIAIAAGIAIACCTAIAACTCGCTCTTCATCCAGACAGACGAAACAAATATTC 141
      |||
753 IleValPheTyrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeuGlnIleVal 772
      |||
122 ATAGTATTCATATACTGCTCTTCATGCGAGATCGAAACAAATATTCGAGATGCT 181
      |||
773 SerThrArgProThrLeuSerProIleProHisIleProArgSerProTyrLysPhe 792
      |||
182 TCCACCTAGCTCTCTTCATGCGAGATCGAAACAAATATTCGAGATGCT 243
      |||
793 ProSerSerProLeuArgIleProGlyGlyAsnIleTyrIleSerProLeuLysSerPhe 812
      |||
242 CAlICIAIAAGIAIACCTAIAACTCGCTCTTCATGCGAGATCGAAACAAATATTC 401
      |||
813 TyrLysIleSerGlnGlyLeuProThrProThrLysMetThrProArgSerArgIleLeu 832
      |||
302 TAlAlTTCACACCCCTCGGAA 361
      |||
833 ValSerIleGlyGlnSerPheGlyThrSerGlnLysPheGlnLysIleAsnGlnMetVal 852
      |||
362 GAlICIAIAAGIAIACCTAIAACTCGCTCTTCATGCGAGATCGAAACAAATATTC 421
      |||
853 CysAsnSerAspArgValLeuLysAsnSerAlaGlnGlnGlnGlnGlnGlnGln 872
      |||
422 TGTAAATGCTCTCTTCATGCGAGATCGAAACAAATATTCGAGATGCT 481
      |||
873 TyrLysIleGlnArgGlnSerPheGlyThrSerGlnLysPheGlnLysIleAsnGlnMetVal 892
      |||
482 AAAAAATGCTCTCTTCATGCGAGATCGAAACAAATATTCGAGATGCT 541
      |||
893 GlyGlnSerLysPheGlnGlnLysLeuAlaGlnMetThrSerThrArgThrArgMetGln 912
      |||
542 GGTACAGTCTCTCTTCATGCGAGATCGAAACAAATATTCGAGATGCT 601
      |||
913 TyrGlnLysMetAsnAspSerMetAspThrSerAsnLysGlnGlnGlnGlnGlnGln 928
      |||
602 AAAGTGAAGAAATACAGCTTAAATTCACCAATTCATGTCACAGCAACACAGATCTCTCT 649

RESULT 9
A1646038 832 bp mRNA linear EST 29-APR 1999
DEFINITION W54410.x1 Soares_mammary_gland_NLMG Mus musculus cDNA clone
IMAGE:1247610 similar to db M25301 Mouse retinoblastoma
susceptibility protein (MOUSE);, mRNA sequence.
A1646038
VERSION A1646038.1 GI:4724513
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 832)
REFERENCE NIH 5541 http://www.ncbi.nlm.nih.gov/ncicqap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: eugene.femail.nih.gov
This clone is available publicly, free through LNCX, contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:561298
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: polyI not found
High quality sequence stop: 482.
Location/Qualifiers
1..832
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1247610"
/clone_lib="Soares_mammary_gland_NLMG"
/sex="female (lactating)"

```









```
tel: +55-11-2704922
fax: +55 11 2707001
Email: asimov@midia.org.br
This sequence was derived from the FAFFSPATITE Human Cancer Genome
Project. This entry can be seen in the following URL
http://www.ncbi.nlm.nih.gov/seq/submit/11185541410941211
Seq primer: puc18 forward
200 014-c06st3-2000-02-21st4-1)
High quality sequence start: 58
High quality sequence stop: 669.
Location/Qualifiers
    1..694
    /organism="Homo sapiens"
    /date="17-Jun-2006"
    /clone_lib="BR0642"
    /pos:1 to "A111"
    /note="organ: breast; Vector: puc18; site_1: Sma1; Site_2:
    Sma1; A mini-library was made by cloning products derived
    from ORFEST PCR (U.S. letters patent application No. 196
    716 - Ludwig Institute for Cancer Research) profiles
    into the puc18 vector. Reverse transcription of tissue
    mRNA and cDNA amplification were performed under low
    stringency conditions."
    211 3 129 3 127 9 226 3
    ; Others
```

Alignment Scores:	
Pred. No.:	694
Score:	1040.00
Percent Similarity:	96.34%
Best local Similarity:	94.95%
Query Match:	21.47%
DB:	10
Length:	694
Matches:	207
Conservatives:	3
Mismatches:	7
Indels:	0
Gaps:	0

US-09 026-459A-2 (1-928) x BE082846 (1-694)

[illegible]

b6		    CTGAAAGTCGTGGTCTTAATCGACGCACCACAATAACACCAATGCCAATGAATAI78  
c9	60*	TTLtLqSsrfoVilAtQsnrFcdLySLyGLfYsoFrThrThrArGal622              
d3	77	GATCCTTGCCCTTTAACACTCATCAAATTAAAAGCTTTACACCGGGCTCA26  
<b>RESULT 14</b>		
AW583181		
LOCUS		
DEFINITION		
aw583181 Human Pancreatic Islets Homo sapiens cDNA clone IMAGE:5638049 5' similar to gb:M5400 RETINOBLASTOMA-ASSOCIATED PROTEIN (HUMAN); gb:M26391 Mouse retinoblastoma susceptibility protein (MOUSE) mRNA sequence.		

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

## REFERENCE

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

FILE  
JOURNAL  
COMMENT

FEATURES  
SOURCE

BASE: COUNT  
ORIGIN

Alignment:  
Pred. No.:  
Score:  
Percent Sim:

Best Local  
Query Match  
DB:

50 122

100



```
QY 555 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
16 556 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
16 557 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
QY 575 1aSerLysAspArgGlyGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeup 595
16 576 1aSerLysAspArgGlyGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeup 595
16 577 1aSerLysAspArgGlyGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeup 595
QY 577 1aSerLysAspArgGlyGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeup 595
16 578 1aSerLysAspArgGlyGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeup 595
QY 595 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
16 596 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
16 597 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
QY 597 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
16 598 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
QY 615 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
16 616 5sAcuUleMeUGlUserLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
QY 645 1aPheGlaThrGlnLysProLeuLys 643
16 646 1aPheGlaThrGlnLysProLeuLys 643
16 647 1aPheGlaThrGlnLysProLeuLys 643
```

Search completed: January 19, 2003, 05:05:07  
Run time : 2072.11 secs

GenCore version 5.1.3  
Copyright (c) 1994 - 2003 CompuGen Ltd.

em nucleotide nucleotide search, using sw model

em on: January 16, 2003, 15:29:22 : Search time 2029.61 Seconds  
(without alignments)  
18469,506 Million cell updates/sec

em on: US 09 026 459A-28

Perfect score: 3455

Sequence: 1 GCGGTCACGACGACACGCG.....AAATGAGGATTATGATACG 3455

Scoring table: IDENTITY NDC

Gapop 10.0 , Gapext 1.0

Searches: 16154066 seqs 806774476 residues

Total number of hits satisfying chosen parameters: 4208132

Minimum hit seq length: 0

Maximum hit seq length: 20000000

Post processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

- Database: EST:
- 1: em\_estb1\*
  - 2: em\_estb1\*
  - 3: em\_estb1\*
  - 4: em\_estb1\*
  - 5: em\_estb1\*
  - 6: em\_estb1\*
  - 7: em\_estb1\*
  - 8: em\_estb1\*
  - 9: qb\_est1\*
  - 10: qb\_est2\*
  - 11: qb\_est3\*
  - 12: qb\_est4\*
  - 13: qb\_est5\*
  - 14: qb\_est6\*
  - 15: em\_estb1\*
  - 16: em\_estb1\*
  - 17: qb\_est1\*
  - 18: em\_estb1\*
  - 19: em\_estb1\*
  - 20: em\_estb1\*
  - 21: em\_estb1\*
  - 22: em\_estb1\*
  - 23: em\_estb1\*
  - 24: em\_estb1\*
  - 25: em\_estb1\*
  - 26: em\_estb1\*
  - 27: em\_estb1\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	929.4	26.9	1024	AK011246	AK011246 Mus musculus
2	910.2	26.3	1024	BM450941	BM450941 AGENT00RT
3	909.4	26.3	1051	BM561083	BM561083 AGENT00RT
4	866	24.8	1081	BQ439221	BQ439221 AGENT00RT
5	844.8	24.5	881	HQ220275	HQ220275 AGENT00RT
6	786	22.7	1001	BM475603	BM475603 AGENT00RT

7	752.4	21.8	1122	13	HM545881
8	726.4	21.0	727	9	AL597611
9	720.6	20.9	880	14	BQ222227
10	699.6	20.2	824	12	BQ610661
11	698.2	20.2	792	9	AL564757
12	692.2	20.0	767	12	BQ576456
13	691.8	20.0	972	13	HM461194
14	678	19.6	835	13	HM454724
15	660	19.1	674	13	B1868459
16	643.6	18.6	797	10	BE549278
17	642.2	18.6	681	12	BQ616219
18	617	17.9	618	13	HM264009
19	614	17.8	694	10	BE082846
20	605.2	17.5	624	10	AW584181
21	605	17.5	605	9	AL599633
22	596.4	17.3	922	12	BQ254543
23	580	16.8	580	9	AL599105
24	575.2	16.6	914	9	AA763411
25	566.3	16.4	588	10	AA763485
26	558	16.2	871	9	AA763485
27	552	16.0	832	9	A1646048
28	530.2	15.3	598	10	AW582967
29	530	15.3	539	9	AL592790
30	528	15.3	572	9	AL125941
31	516.8	15.0	597	12	HE941236
32	513.2	14.9	594	10	AV715533
33	512.2	14.8	734	12	BE144857
34	505.2	14.6	736	13	B1151568
35	493.8	14.3	525	9	AA258255
36	486.4	14.1	584	12	BQ149050
37	484.4	14.0	493	10	AW502887
38	481.4	13.9	484	10	HE168095
39	480.2	13.9	659	13	B1821116
40	476.6	13.8	2083	13	BM454728
41	474.4	13.7	734	13	B1526982
42	472.4	13.7	709	9	A1646097
43	459.2	13.3	497	12	BE943784
44	457	13.2	457	9	A1094215
45	457	13.2	732	9	AL598766

ALIGNMENTS

RESULT 1	AK011246	1304 bp	mkNA	linear	HIC 19 JAN 2002
LOCUS	Mus musculus 10 days embryo whole body cDNA, KIKEN full length				
DEFINITION	embryonic library, clone:260017M22;ref:indiaStoma 1, full length				
ACCESSION	AK011246				
VERSION	AK011246.1	6112847243			
KEYWORDS	HIC; CAP trapper				
SEQUENCE	Mus musculus (strain:5766/66) 10 days embryo cDNA to mRNA, clone:11b81K8N full length enriched mouse cDNA library, clone:260017M22.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kuroki, H., Ohtsuka, Y., Motomatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
REFERENCE	2049374				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kuroki, H., Ohtsuka, Y., Motomatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
REFERENCE	11042159				























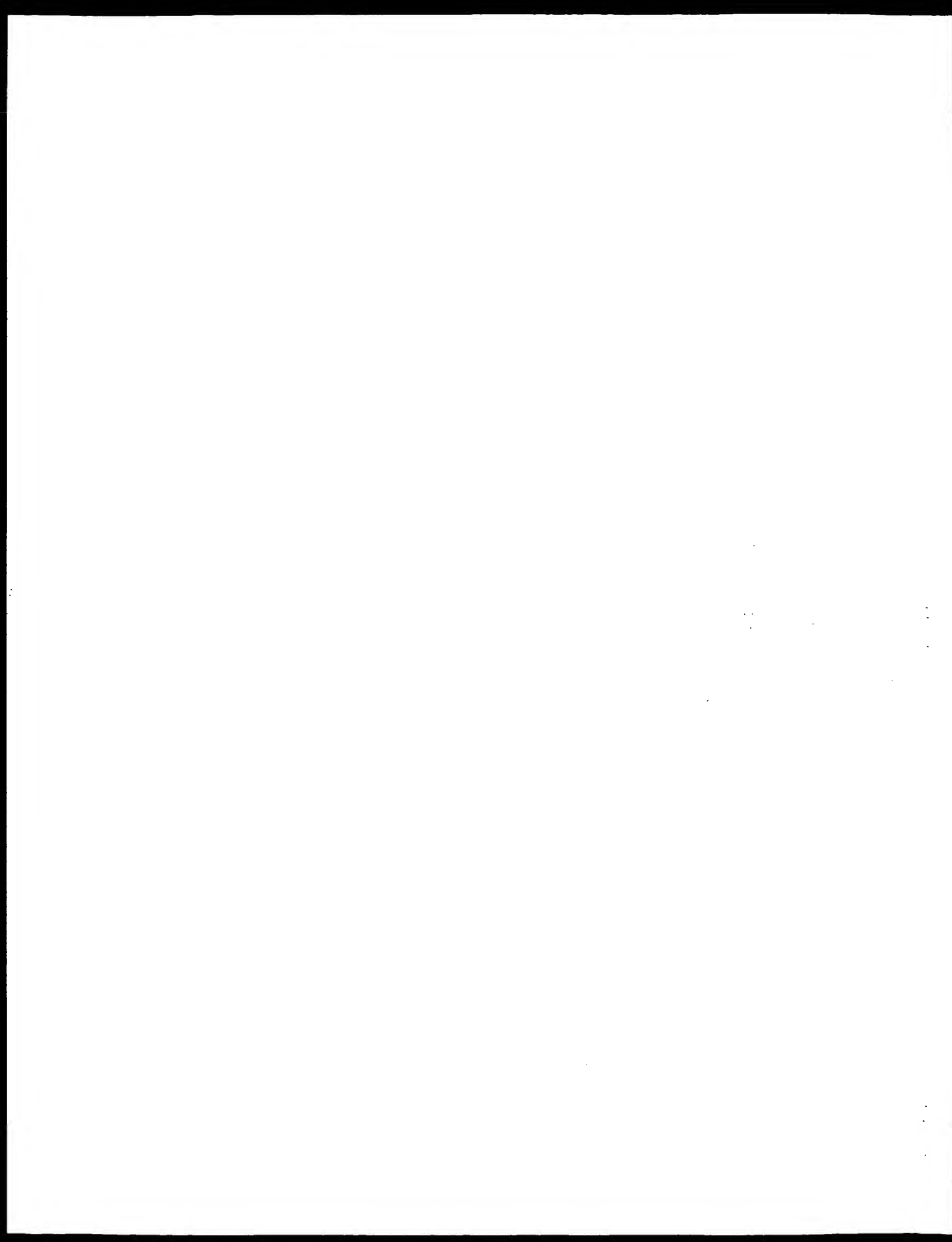




full length clones and constructed by Life Technologies.  
Note: this is a NIH-MSC Library."

BASE COUNT	228 a	124 c	134 g	186 t
ORIGIN				
Query Match.				
Best Local Similarity 99.98; Score 660; DR 14; Length 672;				
Matches 6/17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;				
QY	2321	ATATTCACGCTGGAAGAGTCATATAAAAAATTCACAGAGCTGCTCAACACCAACAAAAA	2380	
DB	1	ATATTCACGCTGGAAGAGTCATATAAAAAATTCACAGAGCTGCTCAACACCAACAAAAA	60	
QY	2381	GCATTCGAAGATCAAGAAATCTTAGTATCAATTCGTGAATCATTCGGGACTTCTGAGAAAT	2440	
DB	61	GCATTCGAAGATCAAGAAATCTTAGTATCAATTCGTGAATCATTCGGGACTTCTGAGAAAT	120	
QY	2441	TCTAGAAAAATTAATCATGATGATATGTATCAACAGACGCTGTGCTCAAAAAAGTGTGAG	2500	
DB	121	TCTAGAAAAATTAATCATGATGATGTATCAACAGAGCTGTGCTCAAAAAAGTGTGAG	180	
QY	2501	CAAGTAAAGCTGCTAAACATCTCAAAAAACATATCTTTCATATTCGAAGATCAACATCAAG	2560	
DB	181	CAAGTAAAGCTGCTAAACATCTCAAAAAACATATCTTTCATATTCGAAGATCAACATCAAG	240	
QY	2561	CATATGGAAGTAAATATCTTTCAGAGAGTGTCAAAATTCACAGCAAAACTGGCAGAAATGA	2620	
DB	241	CATATGGAAGTAAATATCTTTCAGAGAGTGTCAAAATTCACAGCAAAACTGGCAGAAATGA	300	
QY	2621	CTTCTACTGAGACAGATGCAAAAGCAGAAATGAATGATGATGATGATGATGATGATGATG	2680	
DB	301	CTTCTACTGAGACAGATGCAAAAGCAGAAATGAATGATGATGATGATGATGATGATGATG	360	
QY	2681	ATCAAGCAAAAATCAAGATTCACAGACCTTGGTGCACATCTGCTGACATCTGCTGATTCAT	2740	
DB	361	ATCAAGCAAAAATCAAGATTCACAGACCTTGGTGCACATCTGCTGACATCTGCTGATTCAT	420	
QY	2741	GTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2800	
DB	421	GTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	480	
QY	2801	CTTCAAGCTCTTTTGTGATATAAATGCTGCAATGCAATGCTGCTGCTGCTGCTGCTGCTG	2860	
DB	481	CTTCAAGCTCTTTTGTGATATAAATGCTGCAATGCAATGCTGCTGCTGCTGCTGCTGCTG	540	
QY	2861	CATTCGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2919	
DB	541	CATTCGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	600	
QY	2920	CATTTAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2979	
DB	601	CATTTAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	660	
QY	2980	ATATTAAGAAATG 2991		
DB	661	ATATTAAGAAATG 672		

Search completed: January 14, 2003, 04:06:32  
Job time: 4063.69 seconds













DB 843 GGGGCTGGGCTCTTGGAGTGGTAAAGGGCAATATAAG 880  
 RESULT 4  
 LOCUS BM475603 1001 bp mRNA Linear EST 05-FEB-2002  
 DEFINITION AGN0001 6480007 NIH\_MGC\_22 Homo sapiens cDNA clone IMAGE:5575683  
 5' mRNA sequence.  
 ACCESSION BM475603  
 VERSION BM475603.1 GI:18724645  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 1001)  
 AUTHORS NIH MGC cDNA Project, a collaboration of  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT contact: Robert Strausberg, Ph.D.  
 Email: rstraus@nhi.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Sequencing by: Agencourt Bioscience Corporation  
 cDNA distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plates: L14M12426 row: 1 column: 04  
 High quality sequence starts: 17  
 High quality sequence stops: 734.  
 High quality sequence location/qualifiers  
 1..1001  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /feature="IMAGE:5575683"  
 /feature="type="embryonal carcinoma, cell line"  
 /note="organ: testis; Vector: pCMV-Sport6; Site:1: NotI;  
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed;  
 Average insert size 2.5 kb. Library enriched for  
 full length clones and constructed by life technologies.  
 Note: this is a NIH\_MGC library."  
 BASE COUNT 489 a 178 c 186 g 288 t  
 CPGTGIN

Alignment Scores:  
 Pred. No.: 2,026 133 Length: 1001  
 Score: 1339.00 Matches: 281  
 Percent Similarity: 93.87% Conservative: 10  
 Best Local Similarity: 90.65% Mismatches: 10  
 Query Matches: 29.02% Gaps: 9  
 50: 13  
 05 09 026 459a 29 (1 895) x BM475603 (1 1001)

QY 192 LeuLysGluAlaAspThrSerThrLysValAspAsnAlaMetSerAlaLeuLysLys 121  
 DB 193 GGGGCTGGGCTCTTGGAGTGGTAAAGGGCAATATAAG 97  
 QY 122 TyrAspValLeuPheAlaLeuPheSerLysLeuGluAlaThrCysGluLeuLeuTyrLe 141  
 DB 123 GGGGCTGGGCTCTTGGAGTGGTAAAGGGCAATATAAG 97  
 QY 141 uThrInProSerSerSerThrGluLeuAspSerAlaLeuValLeuLysValSe 161  
 DB 142 GGGGCTGGGCTCTTGGAGTGGTAAAGGGCAATATAAG 97  
 QY 161 ThrPheLeuPheLeuAlaLysGlyGluValLeuGluPheLeuGluValLeuLysValSe 181  
 DB 162 GGGGCTGGGCTCTTGGAGTGGTAAAGGGCAATATAAG 97  
 QY 216 ThrGluAlaCATTATTTATAGTAAAGGGAGATTAACAAATGAAATGATTTGGTAT 277  
 DB 217 GGGGCTGGGCTCTTGGAGTGGTAAAGGGCAATATAAG 97  
 QY 181 SerPheGluLeuMetLeuLysValLeuAspLysPheLeuLysLeuSerProProMetLe 201

DB 278 TTCAATTCAGTTAAAGTATGCTGCTGCACTATTTATTAAGATCTAAATCAAGCTT 337  
 QY 201 uLeuLysGluProTyrLysThrAlaValLlePheLeuValLysSerLeuAlaLeuLeu 221  
 DB 338 GCTCAAGCAAGCAATATAAGAGCTGTATACGCTATTAAGCTTAAAGCTTCACTGCAATCAAG 397  
 QY 221 GATGCTCTCTTAAATATGCTATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241  
 DB 398 GCGAGGTCAATACAGAGAGTGTACAGATAGTAAACAAATAGAAAAATGATATAGAAATATAT 457  
 QY 241 cGluValLeuCysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 261  
 DB 458 TGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 517  
 QY 261 SASGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281  
 DB 518 AAAATTTATACCTTTATGAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 577  
 QY 281 TCTAAsnLeuSerLysAlaTyrGluGluLleTyrLeuLysAsnLysAspLysAspLysAlaAla 301  
 DB 578 TCAAAATCTTTCTAAATAGTAAAGAAATATATTTAAAAATATAATATATATATATATATATATAT 647  
 QY 301 GlnPheLeuAspHisAspLysThrLeuGluThrAspSerLleAspSerPheLeuThrC 421  
 DB 638 ATATATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 697  
 QY 321 nArqThrProArqLysSerAsnLeuAspGluGluValAsnValLlePheProHisThr 441  
 DB 698 CAGAACACCACCAAAAAAGTAAACCTTGCAAGACAGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCT 757  
 QY 341 oValArqThrValMetAsnThrLleGluGluLeuMetMetLleLeuAsnSerAlaSerAs 461  
 DB 758 AGTTAGAGCTCTTATGAACTATGAACTATGAACTATGAACTATGAACTATGAACTATGAACTATGAA 817  
 QY 361 PCTnProSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 480  
 DB 818 TCAACCTTCAAAAAATGCAATTTCTCTATTTTAAACAAATGCAAGCTGCAATGCTGCTGCTGCTGCTGCT 877  
 QY 380 uSerLleLeuLysArq ValLysAspLleGlyTyrLle LeuLeuLysLys 496  
 DB 878 AAAGCTATACCGGCAAAACAAAGGTAAGGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAA 947  
 QY 397 --PheAlaLysAlaValGly 402  
 DB 938 TTCTGGCTAAAGGCTAGGG 957

RESUI 5  
 AL597811  
 LOCUS AL597811 727 bp mRNA Linear EST 14 AUG 2001  
 DEFINITION DKFZP3100816\_1 313 (cysnrgm: h1cc2) Homo sapiens cDNA clone  
 DKFZP3100816 5' mRNA sequence.  
 ACCESSION AL597811  
 VERSION AL597811.1 GI:15160502  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 727)  
 AUTHORS Poustka A., Wellentauter, K., Mewes, H.W., Well, H., and Wiemann, S.)  
 JOURNAL EST (Poustka A., Wellentauter, K., Mewes, H.W., Well, H., and Wiemann, S.)  
 COMMENT Unpublished (1999)  
 Contact: Poustka A.J.  
 Department: Leitch  
 Max Planck Institute for Molecular Genetics  
 Ihnestrasse 73, 14195 Berlin, Germany  
 Tel: +49-30-84141624  
 Fax: +49-30-84141128  
 Email: poustka@mpg-berlin.dahlem.mpg.de  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer





```

QY 771 -ileTyrlleSerProLeuLys 777
||||| 333 |||||||
DB 728 GAGCTTAAGATTTCCTACCTTACGAGTTCAGTAATCCGAGGAAAGGAAATATCC 787
||||| 333 |||||||

QY 771 -ileTyrlleSerProLeuLys 777
||||| 333 |||||||
DB 748 TATAATTTTCAACCCCTCGGAA 809
||||| 333 |||||||

RESULT 8
BM450031
DEFINITION A-ENC-001_6403544 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528037
5' mRNA sequence.
Accession BM450031
Version BM450031.1 GI:18459071
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 1023)
NIH MGC http://mgi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Published (1999)
Contact: Robert Strausberg, Ph.D.
Email: eusps@mail.nih.gov
Tissue Procurement: ATCC/DCTD/BIP
cDNA Library Preparation: Life Technologies, Inc
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by Applied Biosystems Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/NIH site
Plate: LLAM12204 row: 3 column: 22
High quality sequence stop: 643.
FEATURES
location/Qualifiers
1..1023
/organism="Homo sapiens"
/ab_xref="taxon:9606"
/cclone="IMAGE:5528037"
/cclone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/vector="pGEM-3513 Vector (CMV-SpeI6, Site 1 Not1;
Site 2: SalI; Cloned unidirectionally. Primer: clone AT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 334 a 211 c 197 g 281 t
ORIGIN
1
Alignment Scores:
Prod. No.: 8.24e-109 Length: 1023
Score: 1111.00 Matches: 216
Percent Similarity: 100.000 Conservations: 0
Best Local Similarity: 100.000 Mismatches: 0
Query Match: 24.08% Indels: 0
Gaps: 0
US 09-026-459A 29 (1-895) x BM450031 (1-1023)
QY 680 LysValLysAsnLysAspLeuLysPheValThrAlaLysAspLeuPro 699
||||| 333 |||||||
DB 2 AAAATGAGAAATATAGCTTAATTTAAATTTATCTTAACTACATACATACATCTCT 61
||||| 333 |||||||
QY 730 HisAlaValGlnGluThrPheLysArgValLeuLysGlnGluLysTyrAspSerIle 719
||||| 333 |||||||
DB 42 CATCTTCTTCAAGAGACATTCACAAAGCTCTTTCATCAACACACACGACATGATCTAT 121
||||| 333 |||||||
QY 729 IleValIleThrAsnSerValPheMetGlnArgLeuLysThrAsnLeuLeuLysTyrAla 739
||||| 333 |||||||
DB 122 ATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 181
||||| 333 |||||||
QY 740 SerThrArgProThrLeuSerProLysPheHisIleProArgSerProTyrLysPhe 759
||||| 333 |||||||

```

```

DB 182 TCCACAGCGCGCCGACCTTCACCAATACCTTCACATTCCTCCAAACCCCTACAAATCT 241
||||| 333 |||||||
QY 760 ProSerSerProLeuArgLleProGlyLysAsnIleTyrlleSerProLeuLysSerPro 779
||||| 333 |||||||
DB 242 CCAAGTTCAGGCTTACGGAATCTCGAGAGGAGCAATCAATATTCACGCGCTGAAGAGTCA 401
||||| 333 |||||||
QY 780 TyrLysIleSerGlnGlyLeuProThrProThrLysMetThrIleArgSerAlaGlnGlu 799
||||| 333 |||||||
DB 302 TATATAATTTTCAAGAGTCTCGCAACACCAACAAAAATGACTCCAGATTCACAAATCTTATA 461
||||| 333 |||||||
QY 403 ValSerIleGlyGlnGlySerGlyThrSerGlnGlyLysIleAsnGlnMetVal 819
||||| 333 |||||||
DB 362 GATCAATTCGTCGAAATCATCTCGGACTCTGACAGATTCACAGAAAATAATTCAGATCGGA 421
||||| 333 |||||||
QY 820 CysAsnSerAspArgValLeuLysAsnSerAlaGlnGlySerAsnProProLysProLys 839
||||| 333 |||||||
DB 422 TGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481
||||| 333 |||||||
QY 940 LysLysIleValPheAspLeuLysGlnGlySerAspLeuLysAspLeuLysLeuPro 869
||||| 333 |||||||
DB 482 AAAAATATAGCTTTATATATATATATATATATATATATATATATATATATATATATAT 541
||||| 333 |||||||
QY 960 GlyGlnSerLysPheGlnGlnGlySerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 879
||||| 333 |||||||
DB 542 GCGACATGCAAAATTCACGACGAAATTCGACAGAAAATGACTTCTACGACACGAAATCGCAA 501
||||| 333 |||||||
QY 980 LysLysIleValPheAsnGlnSerMetAspThrSerAsnLysGlnGlnGlnGlnGlnGln 895
||||| 333 |||||||
DB 502 AACGCAAAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649
||||| 333 |||||||
PFSUIT 9
AA763411/c
LOCUS
DEFINITION 914 bp mRNA linear EST 27 JAN 1998
W545404.1 Soares_mammary_gland_NLMG Mus musculus cDNA clone
IMAGE:1247598 5' similar to gb:M15400 RETINOBLASTOMA-ASSOCIATED
PROTEIN (HUMAN); gb:M2639; Mouse retinoblastoma susceptibility
protein (MOUSE);, mRNA sequence.
Accession AA763411
Version AA763411.1 GI:2813158
KEYWORDS EST
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogamathi; Muridae; Murinae; Mus
1 (bases 1 to 914)
Marra, M., Hillier, L., Allen, M., Bowles, M., Diehrich, N., Dubouque, J.,
Geisel, S., Kucaba, T., Levy, M., Le, M., Martin, J., Mettlin, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, R.,
Theisinger, R., Wyllie, F., Lennon, G., Soares, R., Wilson, R., and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty free through HMI; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:661286
Seq primer: -28ml3 rev2 ET from AmerSham
High quality sequence stop: 422.
location/Qualifiers
1..914
/organism="Mus musculus"
/ab_xref="taxon:10090"
/cclone="IMAGE:1247598"
/cclone_lib="Soares_mammary_gland_NLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
FEATURES
SOURCE

```















Genotype version 5.1.1.4  
Copyright (c) 1994 - 2003 CompuGen Ltd.

AM protein nucleic search, using frame\_plus\_pzn model

Run on: January 17, 2003, 10:27:24 : Search time 1727.82 Seconds  
(without alignments)  
7420.587 Million cell updates/sec

Little: us-09-026-459a-37  
Pattern score: 4017  
Sequence: 1 MSRLKKYLVIPATPSKTFP ...TPMKKPKNNISMTSNEEK 781

Scoring table: 648SUM62  
Xgapop 10.0 : Xgapext 0.5  
Ygapop 10.0 : Ygapext 0.5  
Zgapop 6.0 : Zgapext 7.0  
Delop 6.0 : Delext 7.0

Scanned: 16154096 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 6269842

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0K

Maximum Match 10K

Listing first 45 summaries

Command line parameters:  
Model: frame\_plus\_pzn model -DEV-YLB  
Q2: 2.002, 1.033, 0.500, 0.250, 0.125, 0.062, 0.031, 0.015, 0.007, 0.003, 0.001, 0.0005, 0.0002, 0.0001, 0.00005, 0.00002, 0.00001  
DB EST: qfmt-lastap -SUFFIX-rst -MINMATCH=0.1 -LOWPCL=0 -LOPEXT=0  
UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40 -cdi -LIST=45  
-LOCALIGN 200 -THR\_SCORE-pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE-LOCAL  
-OUTFMT-pto -NUPM-ext -HEADS-PTH-500 -MINFN=0 -MAXFN=200000000  
-NO\_XIPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WALL -LOWPCL=0.5 -DEV-YLB -DEV-YLB  
-WARN TIMEOUT=40 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
YAPOP 10 YGAPEXT=0.5 DELOP=6 DELEXT=7

Database 1:

EST: \*  
1: em estba: \*  
2: em estbmc: \*  
3: em estbmc: \*  
4: em estbmc: \*  
5: em estbmc: \*  
6: em estbmc: \*  
7: em estbmc: \*  
8: em estbmc: \*  
9: qb est 1: \*  
10: qb est 2: \*  
11: qb est 3: \*  
12: qb est 4: \*  
13: qb est 5: \*  
14: qb est 6: \*  
15: qb est 7: \*  
16: qb est 8: \*  
17: qb est 9: \*  
18: em qss\_bmc: \*  
19: em qss\_bmc: \*  
20: em qss\_bmc: \*  
21: em qss\_bmc: \*  
22: em qss\_bmc: \*  
23: em qss\_bmc: \*  
24: em qss\_bmc: \*  
25: em qss\_bmc: \*  
26: em qss\_bmc: \*  
27: em qss\_bmc: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1551	38.6	1051	13	BM561083
2	1448	36.0	881	14	BM220275
3	1289	32.1	1001	13	BM475603
4	1254	31.2	727	9	AL597811
5	1243	30.9	1464	11	AK011246
6	1236.5	30.8	880	14	BM220275
7	1208	30.1	835	13	BM453724
8	1111	27.7	1233	13	BM450041
9	1056.5	26.3	914	9	AA764411
10	1043	26.0	752	9	AL564757
11	1037	25.8	605	9	AL599633
12	1030	25.6	654	10	BM082846
13	1016	25.3	623	10	AW583181
14	1013	25.2	871	9	AA764685
15	1034	25.0	922	12	BM253433
16	978	24.3	588	10	AW368234
17	962	23.9	746	13	BI151568
18	959	23.9	594	10	AV715533
19	955	23.8	584	12	BM149050
20	947.5	23.6	744	12	BM144857
21	935	23.3	539	9	AL692790
22	938	22.6	797	10	BE33278
23	882	22.0	572	9	AL12941
24	879.5	21.9	734	13	BI1526982
25	869.5	21.6	625	12	BM243888
26	858	21.4	597	12	BM141236
27	844	21.0	484	10	BM148095
28	833.5	20.7	646	13	BM490416
29	824	20.3	453	10	AW592887
30	815.5	20.3	522	9	AA072780
31	784	19.5	2083	13	BM456728
32	751	18.7	453	10	AW503514
33	683	17.0	438	12	BM27047
34	681.5	17.0	799	14	CM2304
35	667	16.6	651	13	BI196841
36	659	16.4	480	10	AW591604
37	633	15.8	425	10	BE285796
38	628	15.6	672	13	BI186839
39	626	15.6	521	12	BM466536
40	625.5	15.6	487	12	BM998490
41	613	15.4	415	9	AT745644
42	611	15.2	410	12	BM560103
43	609	15.2	411	9	AT745643
44	593	14.8	580	9	AL599105
45	544.5	13.6	334	12	BM604092

ALIGNMENTS

RESULT 1  
BM561083

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

BM561083 1051 bp mRNA linear EST 20-SEP-2002  
AGENCOURT\_656613 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:5559905  
5' mRNA sequence

BM561083

BM561083.1 GI:18806043

EST

human

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;

1 (bases 1 to 1051)

NIH MGC <http://marc.ncbi.nlm.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

















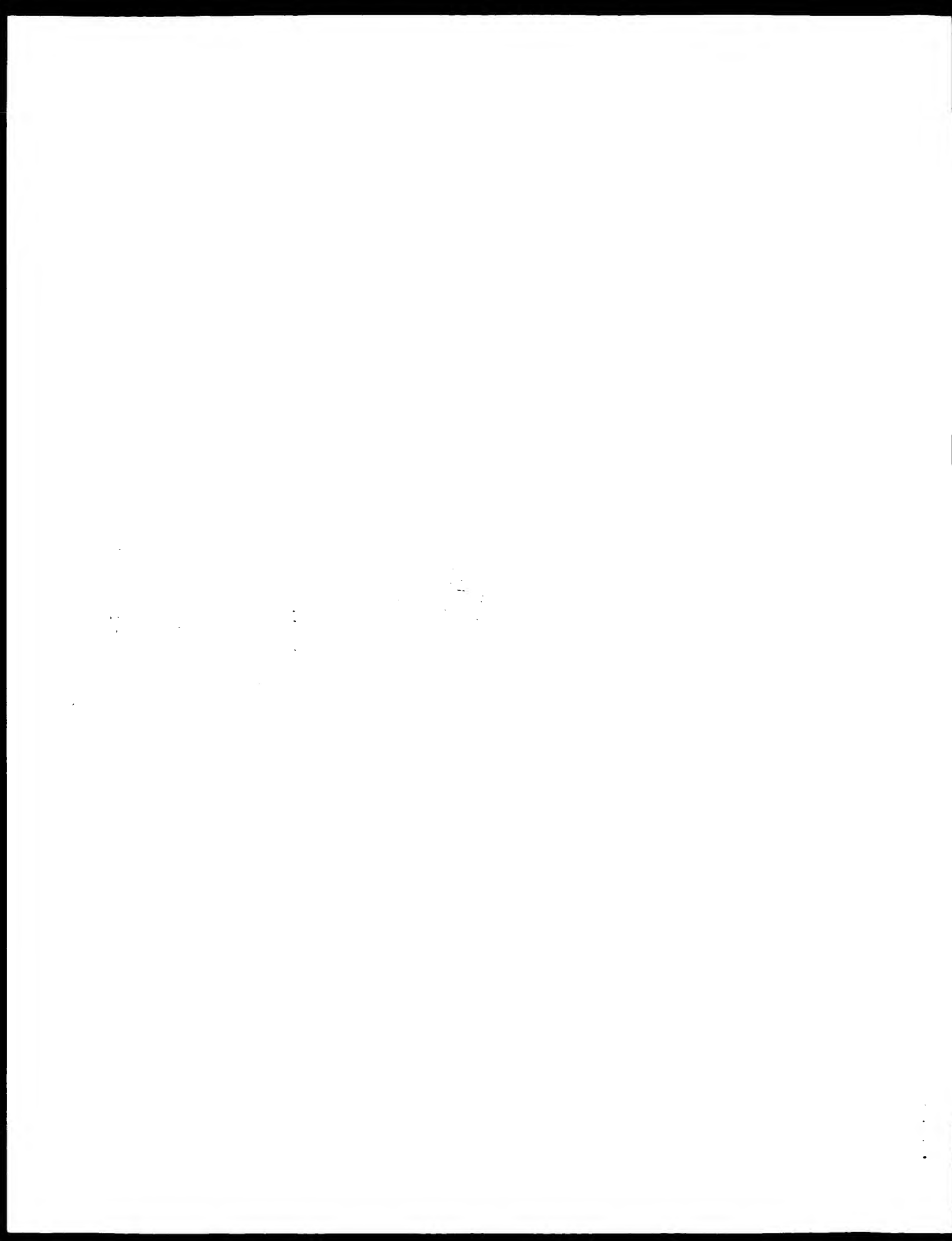














Genature version 5.1.4  
Copyright (c) 1994-2003 CompuGen Ltd.

com nucleotide complete search, using SW model

Run on: January 16, 2003, 15:20:17 : Search time 5160.97 Seconds  
(without alignments)  
17554.275 Million cell updates/sec

Filter: US-09-026-459A-36

Perfect score: 4114

Sequence: 1 GGGGAGTCACAGACGCGG.....AAAACAGCAATATGATAGT 3113

Scoring table:

IDENTITY NDS

Gapop 10.0 : Gapext 1.0

Searched: 2054440 seqs, 14533402878 residues

Total number of hits satisfying chosen parameters: 4106280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenBank: \*  
1: db\_bact: \*  
2: db\_bact: \*  
3: db\_bact: \*  
4: db\_bact: \*  
5: db\_bact: \*  
6: db\_bact: \*  
7: db\_bact: \*  
8: db\_bact: \*  
9: db\_bact: \*  
10: db\_bact: \*  
11: db\_bact: \*  
12: db\_bact: \*  
13: db\_bact: \*  
14: db\_bact: \*  
15: db\_bact: \*  
16: db\_bact: \*  
17: db\_bact: \*  
18: db\_bact: \*  
19: db\_bact: \*  
20: db\_bact: \*  
21: db\_bact: \*  
22: db\_bact: \*  
23: db\_bact: \*  
24: db\_bact: \*  
25: db\_bact: \*  
26: db\_bact: \*  
27: db\_bact: \*  
28: db\_bact: \*  
29: db\_bact: \*  
30: db\_bact: \*  
31: db\_bact: \*  
32: db\_bact: \*  
33: db\_bact: \*  
34: db\_bact: \*  
35: db\_bact: \*  
36: db\_bact: \*  
37: db\_bact: \*  
38: db\_bact: \*  
39: db\_bact: \*  
40: db\_bact: \*  
41: db\_bact: \*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4107	99.8	4600	9	HUMRHARA
2	4107	99.8	4600	9	AX540701
3	4107	99.8	4600	9	HUMRHARA
4	4107	99.8	4600	9	HUMRHARA
5	4105.4	99.8	3242	6	AR022042
6	4105.4	99.8	3242	6	AR022042
7	4105.4	99.8	3242	6	118496
8	4105.4	99.8	3242	6	118497
9	4105.4	99.8	3242	6	105411
10	4099	99.6	4740	9	HUMRHARA
11	4048.4	97.9	4580	6	AX540701
12	4048.4	97.9	4580	6	AX540701
13	4048.4	97.9	4580	6	AX540701
14	4048.4	97.9	4580	6	AX540701
15	4048.4	97.9	4580	6	AX540701
16	4048.4	97.9	4580	6	AX540701
17	4048.4	97.9	4580	6	AX540701
18	4048.4	97.9	4580	6	AX540701
19	4048.4	97.9	4580	6	AX540701
20	4048.4	97.9	4580	6	AX540701
21	4048.4	97.9	4580	6	AX540701
22	4048.4	97.9	4580	6	AX540701
23	4048.4	97.9	4580	6	AX540701
24	4048.4	97.9	4580	6	AX540701
25	4048.4	97.9	4580	6	AX540701
26	4048.4	97.9	4580	6	AX540701
27	4048.4	97.9	4580	6	AX540701
28	4048.4	97.9	4580	6	AX540701
29	4048.4	97.9	4580	6	AX540701
30	4048.4	97.9	4580	6	AX540701
31	4048.4	97.9	4580	6	AX540701
32	4048.4	97.9	4580	6	AX540701
33	4048.4	97.9	4580	6	AX540701
34	4048.4	97.9	4580	6	AX540701
35	4048.4	97.9	4580	6	AX540701
36	4048.4	97.9	4580	6	AX540701
37	4048.4	97.9	4580	6	AX540701
38	4048.4	97.9	4580	6	AX540701
39	4048.4	97.9	4580	6	AX540701
40	4048.4	97.9	4580	6	AX540701
41	4048.4	97.9	4580	6	AX540701

# ALIGNMENTS

Result 1	HUMRHARA	Human retinoblastoma associated (RB1) mRNA, complete cds.	Linear	Pos 12-101, 1995
LOCUS	HUMRHARA	4600 bp		
DEFINITION	Human retinoblastoma associated (RB1) mRNA, complete cds.			
ACCESSION	M34647.1	GI:190945		
VERSION	M34647.1	GI:190945		
KEYWORDS	retinoblastoma protein.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 4600)			
AUTHORS	Friedman, S.H., Borowitz, J.M., Gerber, M.R., Ward, X.F., Beermann, E., Li, F.P., and Weinberg, R.A.			
TITLE	Deletions of a DNA sequence in retinoblastomas and mesenchymal			









variation	<p>179</p> <p>/gene="RB1"</p> <p>/note="Tumor G051q unilateral mutation causing a tentative wrong splice; G00-118-734"</p> <p>/citation="51"</p> <p>/replace="t"</p> <p>275..401</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-2</p> <p>402..517</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-3</p> <p>506..509</p> <p>/gene="RB1"</p> <p>/note="RHF64 bilateral familial deletion mutation causing premature stop"</p> <p>/replace="ta"</p> <p>518..637</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-4</p> <p>537..542</p> <p>/gene="RB1"</p> <p>/note="Tumor G0561 unilateral frameshift mutation causing premature stop"</p> <p>/citation="51"</p> <p>/replace="ta"</p> <p>621..622</p> <p>/gene="RB1"</p> <p>/note="Tumor G0537 bilateral frameshift mutation causing premature stop; G00-118-734"</p> <p>/citation="51"</p> <p>/replace="c"</p> <p>638..676</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-5</p> <p>677..744</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-6</p> <p>745..855</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-7</p> <p>856..944</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-8</p> <p>999..1076</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-9</p> <p>1077..1186</p> <p>/gene="RB1"</p> <p>/note="G00-118-734"</p> <p>/number-10</p> <p>1095</p> <p>/gene="RB1"</p> <p>/note="Tumor G0559 mutation causing premature stop ; G00-118-734"</p> <p>/citation="51"</p> <p>/replace="t"</p>
exon	
exon	
variation	
exon	
variation	
exon	
exon	
variation	
exon	
variation	

106 999 GCGAAAAAGGTTTAAATTTAAAAATTTTAAAGCTTTTAAAGATTTCTTGGACATTTCTTAAACA 1058  
107 487 TCTAATGAGATTCGAGAAAGGTTTCTAAAGAGATAGAAATTAATTTAAATTTAA 546  
108 1059 TCTAAGAGATTCGAGAGGTTTAAAGATTTCTTAAAGATAGAGAAATTTATCTTAA 1118  
109 547 AATTAAGATTCGAGAGGATTTATTTTGGATTCATGATTTCTTAAAGATTTCTTAA 606  
110 1119 AATTAAGATTCGAGAGGATTTATTTTGGATTCATGATTTCTTAAAGATTTCTTAA 1178  
111 607 ATAGAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 666  
112 1179 ATAGAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1238  
113 667 GTAAATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 726  
114 1249 GTAAATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1298  
115 727 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 786  
116 1299 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1358  
117 787 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 846  
118 1459 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1418  
119 847 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 906  
120 1479 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1478  
121 907 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 966  
122 1479 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1538  
123 967 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1026  
124 1539 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1598  
125 1027 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1086  
126 1599 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1658  
127 1087 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1146  
128 1659 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1718  
129 1147 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1206  
130 1719 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1778  
131 1207 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1266  
132 1779 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1838  
133 1267 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1326  
134 1839 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1898  
135 1327 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1386  
136 1899 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1958  
137 1387 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1446  
138 1959 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2018  
139 1447 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1506  
140 2019 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2078  
141 1507 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1566  
142 2079 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2138

106 1567 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1626  
107 2139 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2198  
108 1627 CTGAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1686  
109 2199 CTGAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2258  
110 1687 ATGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1746  
111 2259 ATGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2318  
112 1747 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1806  
113 2419 TATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2478  
114 1807 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1866  
115 2479 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2538  
116 1867 ATGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1926  
117 2439 ATGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2498  
118 1927 AGGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 1986  
119 2499 AGGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2558  
120 1987 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2046  
121 2559 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2618  
122 2047 AGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2106  
123 2619 AGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2678  
124 2107 ATAAATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2166  
125 2679 ATAAATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2738  
126 2167 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2226  
127 2739 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2798  
128 2227 AGTAAATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2286  
129 2799 AGTAAATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2858  
130 2287 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2346  
131 2859 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2918  
132 2347 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2406  
133 2419 AATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2978  
134 2407 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 2466  
135 2479 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 3038  
136 2467 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 3126  
137 3039 GATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 3098  
138 2527 ATGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 3186  
139 3099 ATGATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 3258  
140 2587 AGTAAATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 3646  
141 3159 AGTAAATTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCT 4218

2447 ATGAGGCTTATGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 2706  
DB 4219 ATGAGGCTTATGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3278  
QY 2207 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 2766  
DB 4279 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3338  
QY 2767 AACATGAAACCGCTTACAGGAGGCTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 2826  
DB 4339 AACATGAAACCGCTTACAGGAGGCTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3398  
QY 2827 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 2886  
DB 4399 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3458  
QY 2987 AACATGAAACCGCTTACAGGAGGCTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 2946  
DB 4459 AACATGAAACCGCTTACAGGAGGCTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3518  
QY 2947 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3006  
DB 4519 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3578  
QY 3007 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3066  
DB 4579 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3638  
QY 3113 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3113  
DB 4609 TACATATAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 3695  
RESULT 4  
AR072041 3232 bp DBA Linear PAT 18 FEB 2000  
DEFINITION Sequence 1 from patent US 5912236.  
ACCESSION AR072041  
VERSION AR072041.1 GI:7222919  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1. Hudes 1 to 3232  
AUTHORS Xu, H.-J.; Hu, S.-X. and Benedict, W.F.  
TITLE Broad-spectrum tumor suppressor genes: gene products and methods for tumor suppressor gene therapy  
JOURNAL Patent: US 5912236 A 1 15-JUN-1997.  
FEATURES  
Location/Qualifiers  
SOURCE 1. 3232  
ORIGIN 1086 a 597 c 566 g 983 t  
Quality Match 99.8%; Score 3105.4; DB 6; Length 3232;  
Best Local Similarity 100.0%; Fwd. No. 0;  
Matches 3106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 7 AAGCAGAGCTGTACAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 66  
DB 124 ATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 183  
QY 67 AAGCAGAGCTGTACAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 126  
DB 184 AAGCAGAGCTGTACAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 243  
QY 127 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 186  
DB 244 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 303  
QY 187 AAGCAGAGCTGTACAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 246  
DB 404 AAGCAGAGCTGTACAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 363

QY 247 AAATCTCAGCTCCCATGTTGCTCAAGAGGAGTATAAAAGAGGCTGTATACGATTAAT 306  
DB 364 AAATCTCAGCTCCCATGTTGCTCAAGAGGAGTATAAAAGAGGCTGTATACGATTAAT 423  
QY 307 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 466  
DB 424 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 483  
QY 367 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 426  
DB 484 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 543  
QY 427 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 486  
DB 544 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 603  
QY 487 TATATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 546  
DB 604 TATATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 663  
QY 547 AAATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 606  
DB 664 AAATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 723  
QY 607 ATAGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 666  
DB 724 ATAGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 783  
QY 667 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 726  
DB 784 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 843  
QY 727 AAATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 786  
DB 844 AAATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 903  
QY 787 AAATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 846  
DB 904 AAATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 963  
QY 847 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 906  
DB 964 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1023  
QY 907 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 966  
DB 1024 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1083  
QY 967 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1026  
DB 1084 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1143  
QY 1027 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1086  
DB 1144 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1203  
QY 1087 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1146  
DB 1204 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1263  
QY 1147 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1206  
DB 1264 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1323  
QY 1267 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1266  
DB 1324 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1383  
QY 1267 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1326  
DB 1384 GATGAGGAGTGGTCTCTATACAGGAGGCTCTCTGACTACTCTGCTCTTTTG 1443











QY 306/ TATTTTTCACCAATTTATGTTTTTAAATGAGAGATATGATAGT 4114  
 Db 3184 TATTTTTCACCAATTTATGTTTTTAAATGAGAGATATGATAGT 4230

RESULT 7  
 11847/6 11847/ 11847/ 4242 bp DNA linear PAT 07-OCT-1996  
 DESCRIPTION Sequence 2 from patient US 5496741.  
 VERSION 11847/ 11847/  
 KEYWORDS 11847/1 61:159RH5.2

SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 4242)  
 AUTHORS X.H. J., H.S. X. and Benedict W.F.  
 TITLE Broad spectrum tumor suppressor genes, gene products and methods  
 for tumor suppressor gene therapy  
 JOURNAL Patent US 5496741 A 2 05 MAR 1996  
 FEATURES  
 Location/Qualifiers  
 1..4242  
 /organism="unknown"

BASE COUNT 983 a 566 c 597 g 1086 t  
 GC101N  
 Query Match 99.9% Score 4105.4; DB 6; Length 4242;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 4109; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 66  
 Db 3113 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 3054  
 QY 67 AATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 126  
 Db 3054 AATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2994  
 QY 127 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 186  
 Db 2994 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2934  
 QY 187 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 246  
 Db 2934 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2874  
 QY 247 AATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 306  
 Db 2874 AATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2814  
 QY 307 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 366  
 Db 2814 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2754  
 QY 367 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 426  
 Db 2754 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2694  
 QY 427 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 486  
 Db 2694 GATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2634  
 QY 487 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 546  
 Db 2634 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2574  
 QY 547 AATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 606  
 Db 2574 AATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2514  
 QY 607 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 666  
 Db 2514 ATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2454

QY 667 GAAATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 726  
 Db 2453 GAAATGTCAGATCTGTGGAAGATATGATGATGTTTGGATCTCTTCAGCAATTTGGAAAGG 2494  
 QY 727 ATTTTAAATTCAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786  
 Db 2493 ATTTTAAATTCAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2444  
 QY 787 ACATGCAATTCAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846  
 Db 2433 ACATGCAATTCAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2274  
 QY 847 GAGAAATTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906  
 Db 2273 GAGAAATTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2214  
 QY 907 CTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966  
 Db 2213 CTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2154  
 QY 967 TATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026  
 Db 2153 TATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2094  
 QY 1027 GATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086  
 Db 2094 GATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2044  
 QY 1087 TCTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1146  
 Db 2043 TCTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1974  
 QY 1147 TTTTACAAATTCAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206  
 Db 1973 TTTTACAAATTCAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1914  
 QY 1207 AATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266  
 Db 1913 AATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1854  
 QY 1267 CTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326  
 Db 1853 CTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1794  
 QY 1327 TCTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386  
 Db 1793 TCTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1744  
 QY 1387 TCTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446  
 Db 1743 TCTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1674  
 QY 1447 GATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1506  
 Db 1673 GATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1614  
 QY 1507 TATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1566  
 Db 1613 TATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594  
 QY 1567 GATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626  
 Db 1553 GATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1494  
 QY 1627 CTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1686  
 Db 1493 CTGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1444  
 QY 1687 ATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746  
 Db 1443 ATGTCAGATCTGTGGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1474











1427 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1486  
 1428 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1821  
 1429 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1446  
 1430 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1881  
 1431 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1506  
 1432 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1941  
 1433 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1566  
 1434 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2001  
 1435 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1626  
 1436 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2061  
 1437 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1686  
 1438 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2121  
 1439 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1746  
 1440 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2181  
 1441 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1806  
 1442 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2241  
 1443 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1866  
 1444 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2301  
 1445 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1926  
 1446 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2361  
 1447 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 1986  
 1448 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2421  
 1449 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2046  
 1450 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2481  
 1451 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2106  
 1452 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2541  
 1453 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2166  
 1454 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2601  
 1455 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2226  
 1456 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2661  
 1457 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2286  
 1458 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2721  
 1459 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2346  
 1460 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2781  
 1461 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2406  
 1462 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2841  
 1463 TTTGTTTGGCTTAAATTTCTGTTCTGAGATAATACACTGACAGATATATATATCT 2466

2842 CAGATGAGCTGATATACCTTCAGAGCTGCTGTTATGAGCAATATATATATATATAT 2901  
 2457 CTTTCTGTCGAT 2426  
 2462 CTTTCTGTCGAT 2961  
 2477 ATGTTGATGCTTAT 2406  
 2482 ATGTTGATGCTTAT 3021  
 2487 ATGTTGATGCTTAT 2446  
 2492 ATGTTGATGCTTAT 3081  
 2497 ATGTTGATGCTTAT 2706  
 2502 ATGTTGATGCTTAT 3141  
 2507 ATGTTGATGCTTAT 2766  
 2512 ATGTTGATGCTTAT 3201  
 2517 ATGTTGATGCTTAT 2826  
 2522 ATGTTGATGCTTAT 3261  
 2527 ATGTTGATGCTTAT 2886  
 2532 ATGTTGATGCTTAT 3321  
 2537 ATGTTGATGCTTAT 2946  
 2542 ATGTTGATGCTTAT 3381  
 2547 ATGTTGATGCTTAT 3006  
 2552 ATGTTGATGCTTAT 3441  
 2557 ATGTTGATGCTTAT 3066  
 2562 ATGTTGATGCTTAT 3501  
 2567 ATGTTGATGCTTAT 3114  
 2572 ATGTTGATGCTTAT 3549

RESULT 10  
 HUMRBS  
 LOCUS Human retinoblastoma susceptibility mRNA, complete cds  
 DEFINITION M15400  
 VERSION M15400.1 GI:190958  
 KEYWORDS retinoblastoma susceptibility  
 SOURCE Homo sapiens (clone: RB [1.5]) testis retina cDNA to mRNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo  
 REFERENCE 1 (bases 243 to 4740)  
 AUTHORS Lee, W.H., Bookstein, R., Hong, F., Young, L.J., Shew, J.Y., and Lee, E.Y.  
 TITLE Human retinoblastoma susceptibility gene: cloning, identification, and sequence  
 JOURNAL Science 245 (4794), 1494-1499 (1987)  
 MEDLINE 87149066  
 PUBMED 3823889  
 REFERENCE 2 (bases 1 to 480)  
 AUTHORS Lee, W.H., Bookstein, R., Young, L.J., Lin, C.J., Rosenfeld, M.G., and  
 TITLE Molecular mechanism of retinoblastoma gene inactivation in  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (16), 6017-6021 (1988)  
 MEDLINE 88320473



1566 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1567 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1568 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1569 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1570 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1571 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1572 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1573 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1574 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1575 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1576 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1577 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1578 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1579 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1580 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1581 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1582 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1583 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1584 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1585 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1586 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1587 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1588 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1589 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1590 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1591 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1592 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1593 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1594 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1595 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1596 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1597 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1598 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1599 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA  
 1600 TATGTTTATATAAAAGGATGATGGTATGCTATCTGCGGTAAATATCATTTTGGAA

2587 AGTTTAGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 2646  
 3160 AGTGGAGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3219  
 2647 ATGGGCTGAGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 2706  
 3220 ATGGGCTGAGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3279  
 2707 TACATATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 2766  
 3280 TACATATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3339  
 2767 ACATGAGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 2826  
 3340 AACATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3399  
 2827 CCATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 2886  
 3400 CCATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3459  
 2887 AAAATATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 2946  
 3460 AAAATATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3519  
 2947 GCTTGTATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3006  
 3520 GCTTGTATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3579  
 3007 TGAATATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3066  
 3580 TGAATATATGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3649  
 3067 TATTTCTGAGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3114  
 3640 TATTTCTGAGAGAGATTGTTTCTGCTGCAAAATGATGTTGTTTATGAGATATGAGA 3696

RESULT 11  
 A01444  
 LOCUS H.sapiens DNA for 4.6 kb retinoblastoma gene probe.  
 ACCESSION A01444  
 VERSION A01444.1 GI:345125  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 AUTHORS Fung, Y.K.T., T'Ang, A., Murphree, L.A. and Benedict, W.F.  
 TITLE A method for detecting the predisposition to retinoblastoma and a  
 retinoblastoma gene probe  
 JOURNAL Patent: EP 0293266-A 1 30 Nov 1988;  
 RESEARCH DEVELOPMENT FOUNDATION  
 FEATURES  
 Location/Qualifiers  
 Source 1..4580  
 Zorganism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 1477 a 840 c 811 g 1451 t 1 others  
 ORIGIN

Query Match 97.98; Score 648.4; DB 6; Length 4580;  
 Best Local Similarity 99.28; Pred. No. 0;  
 Matches 4083; Conservative 0; Mismatches 22; Indels 2; Gaps 2;  
 7 ATGTCAAGAGCTGTGCAAGATGATGATGTTGTTGCACTTTCAGCAAAATGCAAGG 66  
 428 ATGTCAAGAGCTGTGCAAGATGATGATGTTGTTGCACTTTCAGCAAAATGCAAGG 487  
 67 ACATGAGAGCTGTGCAAGATGATGATGTTGTTGCACTTTCAGCAAAATGCAAGG 126  
 488 ACATGAGAGCTGTGCAAGATGATGATGTTGTTGCACTTTCAGCAAAATGCAAGG 547











Accession AR247841  
Version AR247841.1 GI:21507687

Source unknown

Organism unknown

Classification

Reference 1 (bases 1 to 2994)

Antelman, D., Gregory, P. J., and Wills, K. N.

Retinoblastoma fusion proteins

Patent: US 6379927 A 30 APR 2002

Journal

Location/Qualifiers

Source

1..2994

/organism="unknown"

BASE 1 2994 974 a 618 c 593 g 809 t

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC

GC













```

Db 1660 TGTGAAGAGATTGCTTCCATGATCTGAAATGCTTAATTTAAAGAGCTTGAT 1719
QY 441 PheThrLysValIleGluSerPheIleLysAlaGluGlyAsnLeuThrArgGluMetIle 400
Db 1720 TTTTAAATATGATGAAGATTTTATCAAGAGCAAGCAACTGACAAAGAAATGATA 1779
QY 441 LysHisLeuAlaArgCysGluHisArgIleMetGluSerLeuAlaIlePheLysSer 420
Db 1780 AAACATTTAGAACGALGIGCAATCGCAALCAAGCAATCCCTGCGCTCCACATCA 1839
QY 421 ProLeuPheAspLeuIleLysGlnSerLysAspArgGluGlyProThrAspHisLeuGlu 440
Db 1840 GTTTTATTGATCTTATTAAATATCAAGAGCAAGCAACTGACAAAGAAATGATA 1899
QY 441 SerAlaCysProLeuAsnLeuProLeuGlnAsnHisIleThrAlaAlaAspMetIleLeu 460
Db 1900 TTGATTTGCTCTTAATCTTCTCCAGATAATACACATGAGAGATATATATATTT 1959
QY 461 SerProValArgSerProLysLysGlySerThrThrArgValAlaAsnSerThrAlaAsn 480
Db 1960 TCTCCGTAAGATCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAT 2019
QY 481 AlaGluThrGluAlaThrSerAlaPheGlnThrGlnLysProLeuLysSerThrSerLeu 500
Db 2020 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2079
QY 501 SerLeuPheLysLysValThrAlaGluAlaLeuAlaGluAsnIleLeuCysGlu 520
Db 2080 TACTGTTTAAATATAAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2139
QY 521 ArgLeuLeuSerGluHisProGluLeuGluHisIleLeuThrLeuPheGlnHisThr 540
Db 2140 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2199
QY 541 LeuHisAsnGluTyrGluLeuMetArgAspArgHisLeuAspGlnIleMetMetCysSer 560
Db 2200 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2259
QY 561 MetIleGlyIleCysLysValLysAsnIleAspIleCysPheLysIleIleValThrAla 580
Db 2260 ATATATGATATATATATATATATATATATATATATATATATATATATATATAT 2319
QY 581 TyrLysAspLeuProHisAlaValGlnGluThrPheLysArgValLeuIleLysGluGlu 600
Db 2320 TACAAGAGATCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2379
QY 601 GluTyrAspSerIleIleValPheTyrAsnSerValPheMetGlnArgLeuLysThrAsn 620
Db 2380 GATATGATCTTATATATATATATATATATATATATATATATATATATATATAT 2439
QY 621 IleGluGlnIleValSerThrArgProProThrLeuSerProIleProHisIleProArg 640
Db 2440 ATTTTATATATATATATATATATATATATATATATATATATATATATATATAT 2499
QY 641 SerProTyrLysPheProSerSerProLeuArgIleProGlyLysAsnIleTyrIleSer 660
Db 2500 AGATTTTATATATATATATATATATATATATATATATATATATATATATATAT 2559
QY 661 ProLeuLysSerProTyrLysIleSerGluGlyLeuProThrProThrLysMetThrPro 680
Db 2560 GCGCTGAAGAGTCAATATAAATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2619
QY 681 ArgSerArgIleGluValSerIleGlyGluSerPheGlyThrSerGluLysPheGlnLys 700
Db 2620 AATATGAAGATCTTATATATATATATATATATATATATATATATATATATATAT 2679
QY 701 IleAsnGlnMetValCysAsnSerAspArgValLeuLysArgSerAlaGluGlySerAsn 720
Db 2680 AATAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2739
QY 721 ProProLysProLeuLysLysLeuArgPheAspIleGluGlySerAspGluAlaAspGly 740

```

```

Db 2740 CCGCTTAAGCAACTGAAAAAATGAGCTTTGATATTGAAGCATCAGATGAAGCAGATGGA 2799
QY 741 SerLysHisLeuThrGlyGluSerLysPheGlnGlnLysLeuAlaGlnMetThrSerThr 760
Db 2800 ACTTAAACATTTGATGAGAGAGTCAAAATTCAGCAGAAACATGCGCAGAAATGATCTTAC 2859
QY 761 ArgThrArgMetGlnLysGlnLysMetAsnAspSerMetAspThrSerAsnLysGluGlu 780
Db 2860 CCAATACCAAAATGCAAAAGCAAAATGCAATGATACCAATGATACCTCAACAAAGAAAG 2919
QY 781 Lys 781
Db 2920 AAA 2922
RESUL 3
AR207831 AP207831 2994 bp cDNA library PAT 20 JUN 2002
DEFINITION Sequence from patient US 6379927.
ACCESSION AR207831
VERSION AR207831.1 GI:21507687
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2994)
AUTHORS Kelleman, D., Gregory, R.J., and Wills, K.N.
TITLE Retinoblastoma fusion proteins.
JOURNAL Patent. US 6379927-A 3 30-APR-2002;
FEATURES Location/Qualifiers
Source
1..2994
/organism="unknown"
BASE COUNT 974 a 618 c 593 g 809 t
ORIGIN
Alignment Scores:
Prod. No.: 4,456,296 Length: 2994
Score: 4017.00 Matches: 781
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-026-459a-37 (1-781) x AR207831 (1-2994)
QY 1 MetSerArgIleuLeuLysLysLysTyrAspValIleuPheAlaIleuPheSerLysLeuGluArg 20
Db 580 ATGTCACAGCTGCTTCAAGCAATATGATGATGATGATGATGATGATGATGATGATGATG 639
QY 21 ThrCysGluLeuIleTyrLeuThrGluProSerSerSerIleSerThrGluIleAsnSer 40
Db 640 ACATGTGAACCTTATATATTTGACACACACCCAGCAGTTCGATATCTACGAAATAAATCT 699
QY 41 AlaLeuValIleuLysValSerIlePheIleThrPheLeuAlaLysGlyGlyValIleuGln 60
Db 700 GATTGATGCTTAAATTTTGTGATATATTTTATTAGCTAATGAGCAAGATTTTAA 759
QY 61 MetGluAspLeuValIleSerPheGlnLeuMetIleuCysValIleuAspLysPheIle 80
Db 760 ATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 819
QY 81 LysLeuSerProProMetLeuLeuLysGluProTyrLysThrAlaValIleProIleAsn 100
Db 820 AAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 879
QY 101 GYSCTPTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 880 GATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 949
QY 121 GluAsnAspThrArgIleIleGluValIleuCysLysLysLysGluHisGluCysAsnIle 140
Db 940 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999
QY 141 ValLysAsnValIlePheLysAspPheIleProPheMetAsnSerLeuGlyLeuValThr 160

```

```

1000 GTCAGAAAGTATTTATTTAAAAATTTTATACCTTTTATGAAATCTGTGGATTTGTAACTA 1059
1010 SGTASGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
1020 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119
1030 ASATGASGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 200
1040 AATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179
1050 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 220
1060 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
1070 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
1080 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1299
1090 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 260
1100 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1359
1110 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
1120 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1419
1130 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
1140 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1479
1150 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
1160 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1539
1170 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340
1180 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599
1190 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
1200 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659
1210 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
1220 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1719
1230 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
1240 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1779
1250 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
1260 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1839
1270 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440
1280 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1899
1290 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 460
1300 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1959
1310 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
1320 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2019
1330 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500
1340 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2079
1350 TTTAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520

```

```

2080 TCACGCTTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 2139
521 ATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
2140 CCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2199
541 LeuGlnAsnGluTyrGlnLeuMetArgAspArgHisLeuAspGlnLeuMetMetCysSer 560
2200 CTCAGCAATCAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2259
561 MetTyrGlnLeuCysSerValLysAsnLeuAspLeuLysPheLysLeuValThrAla 580
2260 ATGATGAGCATATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2319
581 TPTLysAspLeuProHisAlaValGlnGlnThrPheLysArgValLeuLeuLysGlnGln 600
2320 TACAAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2379
601 GlnTyrAspSerThrValPheTyrAspSerValPheMetGlnAlaLeuLysThrAsn 620
2380 GAGTATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2439
621 TLeuGlnTyrAlaSerThrArgProProThrLeuSerProLeuProHisLeuProArg 640
2440 ATTTGAGTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2499
641 SerProTyrLysPheProSerSerProLeuArgTLeuProGlnGlnGlnGlnGlnGlnGln 660
2500 AGCTTTACAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2559
661 ProLeuLysSerProTyrLysLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 680
2560 CCCTCAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2619
681 ArgSerArgTLeuValSerTLeuGlnGlnSerPheGlnGlnGlnGlnGlnGlnGlnGlnGln 700
2620 AGATCAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2679
701 TLeuGlnMetValLysAsnSerAspArgValLeuLysAspArgValLeuLysSerAsn 720
2680 ATAAATAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2739
721 ProProLysProLeuLysLysSerValPheAspTLeuGlnGlnGlnGlnGlnGlnGlnGln 740
2740 CTTCTCAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2799
741 SerLysHisLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 760
2800 AGTAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2859
761 ArgThrArgMetGlnLysGlnLysMetAspAspSerMetAspLeuSerAsnLysGlnGln 780
2860 CGAACAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2919
781 Lys 781
III
Db 2920 AAA 2922

```

## RESULT 4

BD009728

Locus

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

CANJ 1 NC

BD009728 2994 bp INA tissue: FAT 31 JAN 2002  
 Tissue specific expression of retinoblastoma protein.

BD009728 40099728  
 40099728.1 GI:18638101

JP 2001503638 A/2.

unidentified.

unclassified.

1 (bases 1 to 2994)

Antelman, D., Gregory, R.J., and Willis, K.N.

Tissue specific expression of retinoblastoma protein

Patent. JP 2001503638 A 2 21 MAR 2001

CANJ 1 NC

COMMENT OS Unidentified  
 PN JP 2001503638-A/2  
 DF 21-MAR-2001  
 PR 15-NOV-1997 JP 1998522958  
 PR 15-NOV-1996 US 08/751517.14-FEB-1997 US 08/401092 PT  
 DEKIAS ANTELMAN RICHARD J GREGORY, KENNETH N WILLS PC  
 COTR: 24 00785 00-A-18/6/00 A61836/12  
 CC Strandedness: Single;  
 CC Topology: linear;  
 FH Key location/Qualifiers  
 FI Source L: 2994  
 FT Location/Qualifiers  
 Location/Qualifiers  
 L: 2994  
 /db\_xref="taxon:32644"  
 /organism="unidentified"

BASE COUNT 974 a 618 c 593 g 809 t  
 ORIGIN

Alignment Scores:  
 Score: 1450.246 Length: 2994  
 4017.00 Matches: 781  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-026-459A-37 (1-781) x B0009728 (1-2994)

21 MetSerArgLeuLeuLysLysTyr:AspValLeuPheAlaLeuPheSerLysLeuGluArg 20  
 26 ATGTAAAGTCTTAAAGAAATATGATGATATTTCTTGTGATTTCTAGCAAAATTGGAAGG 639  
 27 HrcysGluLeuLeuTyrLeuThrGlnProSerSerSerLeuSerThrGluLeuAsnSer 40  
 34 ATATGAACTTATATATTTATTAACAAACCCAGCTTGATATTTCTGAAATATATCT 699  
 41 AlaLeuValLeuLysValSerTrpIleThrPheLeuLeuAlaLysGlyGluValLeuGln 60  
 700 GTATGAGCTGTAAGAAATTTCTTCTGACATATTTATTATTAAGTGAAGATATACAA 759  
 61 MetGluAspLeuValIleSerPheGlnLeuMetLeuLysValLeuAspTyrPheIle 80  
 760 ATGAAGATGATGCTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 819  
 81 LysLeuSerTrpTrpMetLeuLeuLysGluProGlyTyrLysThrAlaValIleProIleAsn 100  
 920 AAACCTTCAAGCTGCAATGTTGCTAAAGAACCATATAAATAAGCTGTTATACCCATTAT 879  
 101 GlySerProArgThrProArgArgGlyGlnAsnArgSerAlaArgIleAlaLysGlnLeu 120  
 990 GGTTCATATTCGAACACCCAGGCGAGCTCAGAACAGGAGTCCACCATAGCAAAACAACIA 939  
 121 GluAsnAspThrArgIleIleLeuValLeuLysGlyGluGlnCysAsnIleAspGlu 140  
 940 AAAATGATACAAATATTTAACTTCTCTATAAATAAATAAATAAATAAATAAATAAATAA 999  
 141 ValLysAsnValTyrPheLysAsnIleIleProPheMetAsnSerLeuGlyLeuValThr 160  
 1000 GTCAAAATGTTTATTTCAAAAATTTTATACCTTTTATGAAATTTCTTGGACTTGTAA 1059  
 161 SerAsnGlyLeuProGluValIleAsnLeuSerLysArgTyrGluGlnIleTyrLeuLys 180  
 1060 TCTAAATGACTTCCAGAGGTTCAAAATCTTCTTAAACCATACGAAGAAATTTATCTTAAA 1119  
 181 AsnLysAspLeuAspAlaArgLeuPheLeuAspHisAspLysThrLeuGlnThrAspSer 200  
 1120 AATAAAGATCTAATGATATTTTATTTGATCTATGATATAAATACTTCTGACTGATCT 1179  
 201 IleAspSerPheGluThrGlnArgThrProArgLysSerAsnLeuAspGluValAsn 220  
 1180 ATAGATATTTTGAACACAGAGAACACACAGGAAAAAGTAAATTTGATGAAGAGGTGAAT 1239

221 ValIleProIleHisThrProValArgThrValMetAsnThrIleGlnLeuMetMet 240  
 DB 1240 GTAATTCCTCCACACACATCCAGCTGTATGACACACTATCCCAACATTAATGATG 1499  
 241 LeuGluAsnSerAlaSerAspGlnProSerGlnAsnLeuIleSerTyrPheAsnAspCys 260  
 DB 1408 ATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 1459  
 261 ThrValAsnProLysGlnSerIleLeuLysArgValLysAspIleGlyTyrIlePheLys 280  
 DB 1360 ACAGTGAATTTCAAAAGAAAGATATCTCAAAAGAGTGAAGGATATAGGATACATCTTAAA 1419  
 281 GlnLysPheAlaLysAlaValGlyGlnGlyCysValGlnLeuIleLysSerGlnAspTyrLys 400  
 DB 1420 CACAAATTTCTTAAACCTTCTGACATACCTTCTGTCGAAATTTGATACAGGATACATAA 1479  
 301 LeuGlyValArgLeuTyrTyrArgValMetGluSerMetLeuLysSerGluGlnGluArg 420  
 DB 1480 CTTCAGATTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1549  
 321 LeuGlnIleGlnAsnPheSerLysLeuLeuAspAsnIlePheHisMetSerIleGln 540  
 DB 1540 TTATCCATTTCAAAATTTTAAAGCAACTTCTGAAATCACAACATTTTTCATATATGCTTTATG 1599  
 341 AlaCysAlaLeuGluValValMetAlaThrTyrSerArgSerThrSerGlnAsnLeuAsp 460  
 DB 1600 GGTCTGCTCTTCTGAGCTTCTTAAATGCTTCAATATAGCAAGATACATCTCAGATTTGAT 1659  
 361 SerGlyThrAspLeuSerPheProIleLeuAsnValLeuAsnLeuLysAlaPheAsp 480  
 DB 1660 TCTGGAATTAATTTGCTTCTTCTGATTTCTGATTTCTGATTTCTGATTTCTGATTTCTGAT 1719  
 381 PheTyrLysValIleGluSerPheIleLysAlaGlnGlyAsnLeuThrArgGlnMetIle 400  
 DB 1720 TTTTCAAAAGCTTCTTCTGAAAGCTTTTATCAAGCTCACAAGCTTCTGATCAAGCAAAATGATA 1779  
 401 LysHisLeuGluArgCysGluHisArgIleMetGluSerLeuAlaIleTrpLeuSerAspSer 420  
 DB 1780 AAACATTTTGAACGATCTGCAACATCTGCAATCTGCAATCTGCAATCTGCAATCTGCAATCTG 1819  
 421 ProLeuPheAspLeuIleLysGlnSerLysAspArgGlnGlyProThrAspHisLeuGlu 440  
 DB 1840 CTTTTATTTGATTTCTTATTTAAATTAATTAAGTACCTTAATTAAGTACCTTAATTAAGTACCT 1899  
 441 SerAlaCysProLeuAsnLeuProLeuGlnAsnHisThrAlaAlaAspMetTyrLeu 460  
 DB 1900 TCTGCTTCTCTCTTCTTCTTCT 1959  
 461 SerProValArgSerProLysLysLysSerThrThrArgValAsnSerThrAlaAsn 480  
 DB 1960 TCTCTCTTAAATCTCTCAAGAAAAAGGTTCAACTACGCTGTGTAATTTCTACTTCTAAAT 2019  
 481 AlaGluThrGlnAlaIleSerAlaPheGlnThrGlnLysProLeuLysSerThrSerLeu 500  
 DB 2020 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2079  
 501 SerLeuPheTyrLysValTyrArgGlnAlaTyrLeuArgLeuAsnThrLysGlnGlu 520  
 DB 2080 TCACTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2139  
 521 ArgLeuLeuSerGluHisProGluLeuGluHisIleIleThrThrLeuPheGluHisThr 540  
 DB 2140 GGTCT 2199  
 541 LeuGlnAsnGluTyrGluLeuMetArgAspArgHisLeuAspGlnIleMetMetCysSer 560  
 DB 2200 CTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2259  
 561 MetTyrGlyIleCysLysValLysAsnIleAspLeuLysThrLysIleIleValIleAla 580  
 DB 2260 ATGATGTCATATTCGAAAGTGAAGATATAGACCTTAAATTTCAAAATTCATTTGATGAACGCA 2319  
 581 TyrLysAspLeuProIleAlaValGlnIleThrPheLysArgValLeuIleLysGluGlu 600









QY	241	116ASG10mMetValProGlySerSerAspAspValLeuLeuGlySerAlaLeuGlySerAlaSo	720
DB	2680	ATATAATCATATGATATATTAACAGCAAGTCTGTCTTAAAAAAGAGTGTCTGAAAGAAAT	2739
QY	242	116GPTG10SPTG10G10S116GATG10TAS116TAS116TAS116TAS116TAS116TAS	745
DB	2740	CTCTCTAAACACATGAAAAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2799
QY	741	SEPLYS116LEU16G10G10G10G10G10G10G10G10G10G10G10G10G10G10G10G10	760
DB	2800	AGTAAATATCTCCAGGACATCT	2859
QY	243	ArgThrArgMetGlnLysGlnLysMetAspSerMetAspSerMetAspSerMetAspSerMet	780
DB	2860	CGAAATGAAATGAAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA	2919
QY	743	116	
DB	2920	AAA 2922	
PEPBLD	7	AP072031	3232 bp DNA
LOCUS	AP072031	Sequence 1 from patent US 5912236.	11-03-97 PAT 18 FEB-2000
DEFINITION	AP072031		
ACCESSION	AP072031		
VERSION	AP072031.1	GI:7222919	
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unclassified.	
REFERENCE		1 (bases 1 to 3232)	
AUTHORS		Xu, H., Li, H., Su, S., X. and Benedict, W. F.	
TITLE		Broad-spectrum tumor suppressor genes gene products and methods for	
JOURNAL		tumor suppressor gene therapy	
FEATURES		Location/Qualifiers	
source		1..3232	
BASE COUNT		1086 a 597 c 546 g 983 t	
ORIGIN		/organism="unknown"	
Alignment Scores:			
Method: BLOSUM62	4 445 246	Length: 3232	
Score:	4017.00	Matches: 781	
Percent Similarity:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	100.00%	Indels: 0	
	6	Gaps: 0	
MS-09 x26.459A-37 (1-781) x AP072031 (1-3232)			
20	1	MetGluAspLeuValIleSerPheGlnLeuMetLeuGlySerValLeuPheAlaLeuPheSerLysLeuGlnArg	20
DB	24	AGTCAAAATCTGTTCAACAGCAATGATGATGTTTGCATCTCTTCCAGCAAAATGCAAAAC	183
QY	21	ThrCysGlnLeuIleTyLeuThrGlnProSerSerIleSerThrGlnIleAsnSer	40
DB	184	ATATGAACTATATATATTTGACACACGACGAGTGCATATCTACTCAATAAATTTCT	243
QY	43	AlaLeuValLeuLysValSerTrpIleTrpPheLeuLeuAlaLysGlyGlnValLeuGln	60
DB	244	GGATTGGTCTTAAAGCTTTCTTCAATATATTTTATAGCTTAAAGGGAGATTTACAA	303
QY	61	MetGluAspLeuValIleSerPheGlnLeuMetLeuCysValLeuAspTrpPheIle	80
DB	404	ATGTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	363
QY	81	LysLeuSerProProMetLeuLeuLysGlnProTyLeuThrAlaValIleProGlnAsn	100
DB	404	AANCTCTACCTCCATGTTGTTCAAGAAATATATTAATTAATTAATTAATTAATTAATTAAT	423
QY	101	GlySerProArgTrpArgArgGlyGlnAsnArgSerAlaArgGlnAlaLysGlnLeu	120





21 ThrCysGluLeuIleTyrLeuThrGlnProSerSerIleSerThrGluIleAsnSer 40  
 184 AATGCGAACTTAAATAATATGAGCAAGGACAGTACATGATATCTACAAATAATTTCT 243  
 41 AlanValLeuLeuLysValSerTrpIleThrPheLeuLeuAlaLysGlyValLeuGln 60  
 243 GATGGTGGTAAAGGHTTGGATACATTTTATAGCTAAAGGAAAGGATATACAA 303  
 61 MetGluAspLeuValIleSerPheGlnLeuMetLeuCysValLeuAspTyrPheIle 80  
 604 ATGAAACATGATCGTGGATATATATTCAGTAAAGCTATAGTGGCTTCACATATTATT 363  
 81 LysLeuSerProProMetLeuLeuLysGlnProTyrLysThrAlaValIleProIleAsn 100  
 664 AAAATGATGCGGATGCTGGTAAAGGAAAGCAATAAAAGAGCTGTATATAGCAATTAAT 423  
 101 GlySerProArgThrProArgGlyGlnAsnArgSerAlaArgAlaAlaLysGlnLeu 120  
 424 GGTTATCTGGAAATGCTAG 483  
 121 GluAspAspThrArgIleIleLeuValLeuGlyCysGlnHisGlnCysAsnIleAspGln 140  
 484 GAAAAATCATACAAATATATGAACTTCCTGAAAGGAAAGCAATGAAATAGATGAG 543  
 141 ValLysAsnValTyrPheLysAsnPheIleProPheMetAsnSerLeuGlyLeuValThr 160  
 544 GTCAAAAATGTTATATCAAAAAATTTATACCTTTTATGAAATTCCTCTGGACATCTAACA 603  
 161 SerAsnGlyLeuProLeuValLeuAsnLeuSerLysArgTyrGlnGlnIleTyrLeuLys 180  
 604 TCTAATGGATCTGCAAGAGTTCAAAATCTTCTAAAGGATAGAGAAATTTATGTTAAA 663  
 181 AsnLysAspLeuAspAlaArgIleGlnPheLeuAspHisSerLysThrLeuLeuThrAspSer 200  
 664 AATAAAGATCTACATGCAAGATATATTTGGATCATGATAAAACCTTCACAGACTGATCT 723  
 201 IleAspSerPheGlnThrGlnArgThrProArgLysSerAsnLeuLeuAspGlyValAsn 220  
 724 ATACAGATTTTGAAGATACAGACACACCCGAAAGCAATGACTTGAATGAAAGAGTGAT 783  
 221 ValIleProProHisThrProValArgThrValMetAsnThrIleGlnGlnLeuMetMet 240  
 784 GTAAATCTGTGCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 843  
 241 IleLeuAsnSerAlaSerAspGlnProSerGlnAsnLeuIleSerTyrPheAsnAsnCys 260  
 844 ATTTAAATTCAGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 903  
 261 ThrValAsnProGlyGlnSerIleLeuLysArgValLysAspIleGlyTyrIlePheLys 280  
 904 ACATGAAATGCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 963  
 281 GluLysPheAlaLysAlaValGlyGlnGlyCysValGlnIleGlySerGlnArgTyrLys 300  
 964 GAAATTTGCTAAAGTCTGGAGAGGGTCTGGTGAATTTGGATTTGATTTGATTTGATTTG 1023  
 301 LeuLysValArgLeuLysTyrArgValMetGlnSerMetLeuLysSerGlnGluArg 320  
 1024 CTGAGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1083  
 321 LeuSerIleGlnAsnPheSerLysLeuLeuAsnAspAsnIlePheHisMetSerLeuLeu 340  
 1084 TAAATGATGAAATGATGAAATGATGAAATGATGAAATGATGAAATGATGAAATGATGAA 1143  
 341 AlaGlyAlaLeuGlnValValMetAlaThrTyrSerArgSerThrSerGlnAsnLeuAsp 360  
 1144 GGTGGATGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1203  
 361 SerGlyThrAspLeuSerPheProTrpIleLeuAsnValLeuAsnLeuLysAlaPheAsp 380  
 1204 TGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1263  
 381 PheTyrLysValIleIleGlnSerPheIleLysAlaGluGlyAsnLeuThrArgGluMetIle 400

1264 TTTTACAAAGTGTATGAAATTTTATCAAGGACAGAGGATTTGACAGAGCAAAATGATA 1323  
 401 LysHisLeuGluArgCysGlnHisArgIleMetGlnSerLeuAlaIlePheLeuSerAspSer 420  
 1324 AAACATTTAGAAATGATGTGAATATGAAATATGGAATATTTGATGAGTATATGAAATGCA 1383  
 421 ProLeuPheAspLeuIleLysGlnSerLysAspArgGluGlyProThrAspHisLeuGln 440  
 1384 CCTTATTTCATCTTATATAACAATCAAGAGACCCAGAGAGCAATCATCATCATTTGAA 1443  
 441 SerAlaCysProLeuAsnLeuProLeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeu 460  
 1444 TCTGCTGTGCTTAAATCTTCTCTCCAGATAATATCAATGCAATATATGATATTT 1503  
 461 SerProValArgSerProLysLysLysGlySerThrThrArgValAsnSerThrAlaAsn 480  
 1504 TCTGCTGTGATGATCTGCAAGGAAAAAGGTTCAATATAGCTGTAAATTTATATGTAAT 1563  
 481 AlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLysProLeuLysSerThrSerLeu 500  
 1564 GCAGACACACACCAAGCTACAGCTTCCAGACCCACAAATGAAATGAAATGAAATGAA 1623  
 501 SerLeuPheTyrLysLysValTyrArgLeuAlaTyrLeuArgLeuAsnThrLeuArgLeu 520  
 1624 TCAGCTGTTTATAAAAAGTGTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 1683  
 521 ArgLeuLeuSerGlnHisProGlnLeuGlnHisIleIleThrThrLeuThrLeuThrHisThr 540  
 1684 CGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1743  
 541 LeuGluAsnGluTyrGlnLeuMetArgAspArgHisLeuAspGlnIleMetMetCysSer 560  
 1744 CTGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1803  
 561 MetTyrGlyIleCysLysValLysAsnIleAspLeuLysPheLysIleIleValThrAla 580  
 1804 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1863  
 581 TyrLysAspLeuProHisAlaValGlnIleThrPheLysArgValLeuIleLysGlnGln 600  
 1864 TATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1923  
 601 GluTyrAspSerIleIleValPheTyrAsnSerValPheMetGlnArgLeuLysThrAsn 620  
 1924 GAGTATCATCTAT 1983  
 621 IleLeuGlnIleValAlaSerThrArgProProThrLeuSerProIleProHisIleProArg 640  
 1984 ATTTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2043  
 641 SerProTyrLysPheProSerSerProLeuArgIleProGlyCysAsnIleTyrIleSer 660  
 2044 AGGCTTATCAAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103  
 661 ProLeuLysSerProTyrLysIleSerGlnGlyLeuProThrProThrLysMetThrPro 680  
 2104 CCCCAGACAGTCCATATAAAATTTTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2163  
 681 ArgSerArgIleLeuValSerIleGlyGlnSerPheGlyThrSerGluLysPheGlnLys 700  
 2164 AGATCAAGATCTTAT 2223  
 701 IleAsnGlnMetValCysAsnSerAspArgValIleLysLysArgSerAlaThrLysSerAsn 720  
 2224 ATAAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2283  
 721 ProProLysProLeuLysLysLeuArgPheAspIleLeuLysSerAspAlaAspLys 740  
 2284 CTTCCCTAAACCTCAAAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2343  
 741 SerLysHisLeuProGlyLysSerLysPheGlnGlnLysLeuAlaLeuMetThrSerThr 760

```

10 2444 AATAAATATGCTGACACACATGCAAAATTCAGCAGCAAAATGCGAGAAATGACATCTACT 2403
11 361 AcuThrAuuMeuGluLysGluLysMetAsnAspSerMetAspThrSerAsnGlyGluGln 780
12 2474 GAAATACAAATGCAAAAGCAGAAATGAATGATGATGATATATTTAAATGACAGAG 2462
13 781 Lys 781
14 2464 AAA 2466
RESULT 10
LOCUS 118497
DEFINITION Sequence 2 from patent US 5496731.
ACCESSION 118497
VERSION 1
FEATURES
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 3232)
AUTHORS Xu, H., Li, B., S., X., and Benedict, W. F.
TITLE Broad spectrum tumor suppressor genes, gene products, and methods
for tumor suppressor gene therapy
PUBLISHED 5496731 A 2005 MAR 19
LOCATION/Qualifiers
1..3232
/organism="unknown"
BASE COUNT 983 a 566 c 597 g 1086 t
ORIGIN
Aldomol Scores:
Prod. No. Length: 3232
Scores: 4017.00 Matches: 781
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US 09-026-459A-37 (1-781) x 118497 (1-3232)
27 1 MetSerArgLeuLeuLysLysThrAspValLeuPheAlaLeuPheSerLysLeuCysArg 20
18 433 AAGCAAAATGCTGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 3054
27 21 ThrLysGluLeuLeuLysLeuThrProSerSerLysLeuLeuLeuLeuLeuLeu 40
18 453 ACAAGCAAAATGCTGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 2994
27 41 AlaLeuValLeuLysValSerTrpPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
18 294 GATGCTGCTGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 2934
27 61 MetGluAspLeuValLysSerPheGlnLeuMetLeuCysValLeuAspTyrPheLeu 80
18 2433 AAGCAAAATGCTGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 2874
27 81 LysLeuSerProProMetLeuLeuLysGluProTyrLysThrAlaValLysProLeuAsn 100
18 2873 AAGCAAAATGCTGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 2814
27 101 GlySerProArgThrProArgArgGlnAsnArgSerAlaArgLeuAlaLysGlnLeu 120
18 2413 GATGCTGCTGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 2754
27 121 GlnAspThrArgGlnLeuValLeuCysLysGluLysGluCysAsnLeuAspGlu 140
18 2753 GAAATGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 2694
27 141 ValLysAsnValTyrPheLysAsnPheLeuProPheMetAsnSerLeuGluLeuValThr 160
18 2753 GCAAAATGCTGCAACCAAGCAAGCAATGCTGACCTTCAGCAAAATGCAAGG 2634

```

```

161 SerAsnGlyLeuProGluValGluAsnLeuSerLysArgTyrGluGluLeuLysLeuLys 180
2633 TTTAAATGCAATTTTAAAGTTGAAAAATCTTTCTAAAGCATAGAGAGAAATTTATCTTAAA 2574
181 AsnLysAspLeuAspAlaArgLeuPheLeuAspHisAspLysThrLeuGlnThrAspSer 200
2573 AATTAAGATGATGATGCAAGATATTTTGGATGATGATGATGATGATGATGATGATGATGAT 2514
201 IleAspSerPheGluThrGluArgThrProArgLysSerAsnLeuAspGluGluValAsn 220
2513 ATAGACAGCTTTTGAACACACAGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 2454
221 ValLeuProHisThrProValArgThrValMetAspThrLysGluGlnCysMet 240
2453 GTAAATGCTGCAACCAAGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2494
241 LeuAsnSerAlaSerAspGlnProSerGlnAsnLeuLysSerTyrPheAsnAspLys 260
2393 ATTTAAATGCAATTTTGAACACACAGAGAAATGATGATGATGATGATGATGATGATGATGAT 2344
261 ThrValAsnProLysGlnSerLeuLeuLysArgValLysAspLeuGlyTyrPhePheLys 280
2333 AAGTAAATGCAATTTTGAACACACAGAGAAATGATGATGATGATGATGATGATGATGATGAT 2274
281 CysLysPheAlaLysValGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 300
2273 GAGAAATGCTGCAACCAAGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2214
301 LeuGlyValArgLeuTyrArgValMetGluSerMetLeuLysSerGluGlnCysArg 320
2213 CTGCAAGCTGCTGCAACCAAGCAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 2154
321 LeuSerLysGlnAspPheSerLysLeuLeuLeuAspAspGlnLeuPheHisMetSerLeuLeu 340
2153 TTATCCATGCAATTTTGAACACACAGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 2094
341 AlaCysAlaLeuGluValValMetAlaThrTyrSerArgSerThrArgThrAsnLeuAsp 360
2093 GGTGCGGTCTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2044
361 SerGlyThrAspLeuSerPheProTrpPheLeuAsnValLeuAsnLeuLysAlaPheAsp 480
2033 TCTGCAACCAATGCTGCTGCAACCAAGCAAGCAATGATGATGATGATGATGATGATGATGATGAT 1974
381 PheTyrLysValLysLeuSerPheLeuLysAlaGluGlyAsnLeuThrArgGluMetLeu 400
1973 TTTTCAAAATGCTGCAACCAAGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1914
401 LysHisLeuLeuArgCysGluHisArgGlnMetGluSerLeuAlaTrpLeuSerAspSer 420
1913 AATCATTTTGAACCAATGCTGCTGCAACCAAGCAAGCAATGATGATGATGATGATGATGATGATGAT 1854
421 ProLeuPheAspLeuLysGlnSerLysAspArgGluGlyProThrAspHisLeuGlu 440
1853 GTTTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1794
441 SerAlaCysProLeuAsnLeuProLeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeu 460
1793 TCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1744
461 SerProValArgSerProLysLysGlySerThrThrArgValAsnSerThrAlaAsn 480
1733 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1674
481 AlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLysProLeuLysSerThrSerLeu 500
1673 GCAGACACACCAACCAACCTGACCTTCCAGAGCCAGAGAGCCATTCGAAATGATGCTGCTGCTGCT 1614
501 SerLeuPheTyrLysValTyrArgLeuAlaTyrLeuArgLeuAsnThrLeuCysGln 520
1613 TCACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1564
521 ArgLeuLeuSerGluHisProGluGluHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 540

```













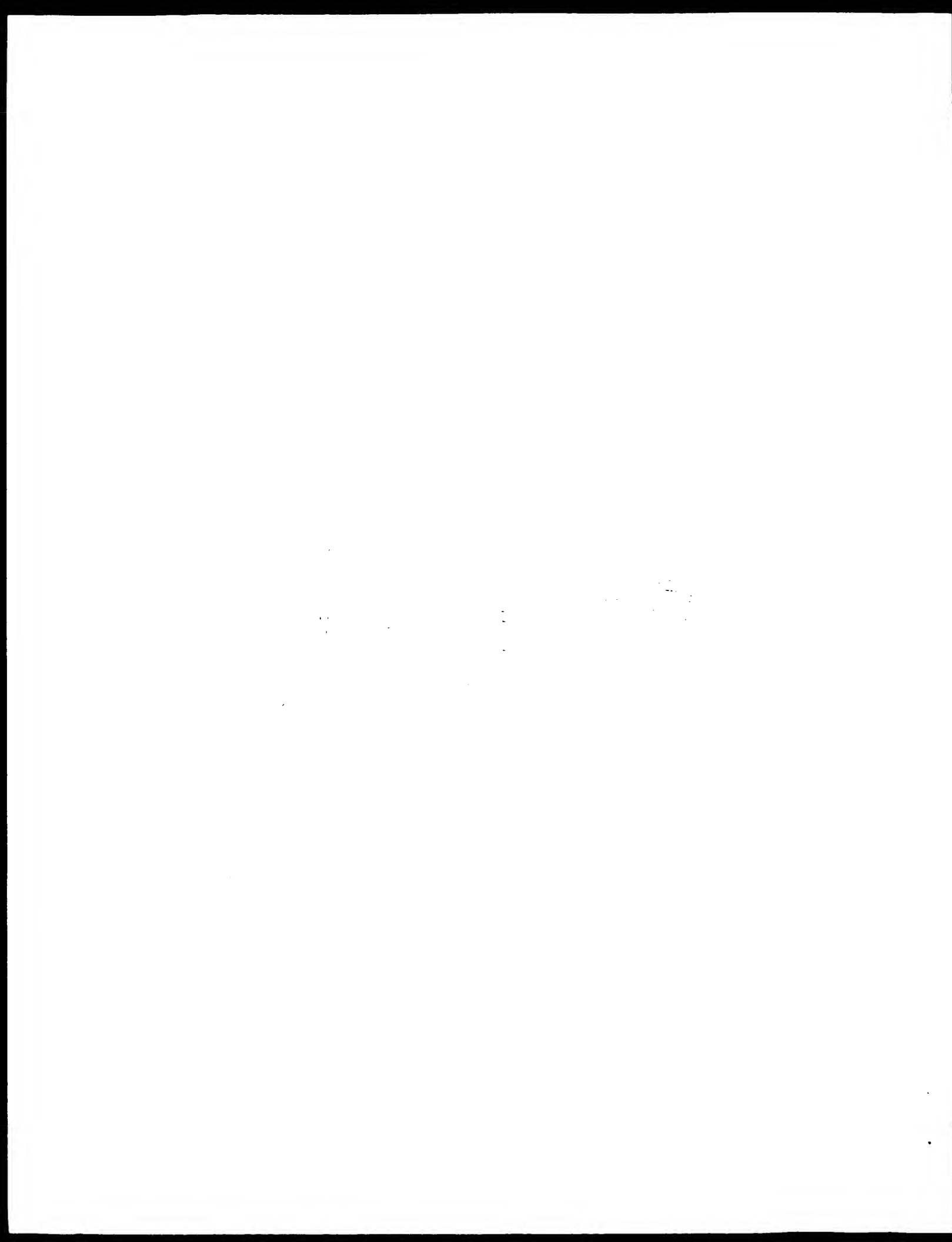












Genome version 5.1.3  
Copyright (c) 1993-2003 CompuGen Ltd.

cm nucleotide nucleotide search, using sw model

Run on: January 16, 2003, 15:20:17 : Search time 5509.12 Seconds  
(without alignments)  
17584.275 Million cell updates/sec

Hit: US 09 026 459A 32

Perfect score: 3323  
Sequence: 1 GGTATATGAGGAGAAAGTTC.....AAATGAGGATATGATAGT 3323

Scoring table:

IDENTITY\_NDC

Gapop 10.0 : Gapext 1.0

Searches: 2054630 seqs, 1455149279 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post processing: Minimum Match 100%

Maximum Match 100%  
Listing first 45 summaries

Database:

1: db\_ba:\*

2: db\_ba:\*

3: db\_ba:\*

4: db\_ba:\*

5: db\_ba:\*

6: db\_ba:\*

7: db\_ba:\*

8: db\_ba:\*

9: db\_ba:\*

10: db\_ba:\*

11: db\_ba:\*

12: db\_ba:\*

13: db\_ba:\*

14: db\_ba:\*

15: db\_ba:\*

16: db\_ba:\*

17: db\_ba:\*

18: db\_ba:\*

19: db\_ba:\*

20: db\_ba:\*

21: db\_ba:\*

22: db\_ba:\*

23: db\_ba:\*

24: db\_ba:\*

25: db\_ba:\*

26: db\_ba:\*

27: db\_ba:\*

28: db\_ba:\*

29: db\_ba:\*

30: db\_ba:\*

31: db\_ba:\*

32: db\_ba:\*

33: db\_ba:\*

34: db\_ba:\*

35: db\_ba:\*

36: db\_ba:\*

37: db\_ba:\*

38: db\_ba:\*

39: db\_ba:\*

40: db\_ba:\*

41: db\_ba:\*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3315	99.8	4600	9	HUMRAH1RA	M33647 Human retin
2	3315	99.8	4600	9	HUMRAH1RA	M33647 Human retin
3	3315	99.8	4600	9	HUMRAH1RA	M33647 Human retin
4	3315	99.8	4600	9	HUMRAH1RA	M33647 Human retin
5	3315	99.8	4600	9	HUMRAH1RA	M33647 Human retin
6	3280.8	98.7	4740	9	HUMRAH1RA	M33647 Human retin
7	3254.8	97.9	4580	6	A01444	M33647 Human retin
8	3222.4	97.0	4232	6	A01444	M33647 Human retin
9	3222.4	97.0	4232	6	A01444	M33647 Human retin
10	3222.4	97.0	4232	6	A01444	M33647 Human retin
11	3222.4	97.0	4232	6	A01444	M33647 Human retin
12	3222.4	97.0	4232	6	A01444	M33647 Human retin
13	2623	78.9	2994	6	AK09189	AK09189 Sequence
14	2623	78.9	2994	6	AK09189	AK09189 Sequence
15	2623	78.9	2994	6	AK09189	AK09189 Sequence
16	2623	78.9	2994	6	AK09189	AK09189 Sequence
17	2623	78.9	2994	6	AK09189	AK09189 Sequence
18	2618.4	78.8	2995	6	AK09189	AK09189 Sequence
19	2281.6	68.8	4591	10	M0303	M0303 cDNA encodi
20	2281.6	68.8	4591	10	M0303	M0303 cDNA encodi
21	2273.4	68.4	4432	10	M0303	M0303 cDNA encodi
22	1473.2	44.3	3456	5	000113	000113 Gallus gall
23	1465	44.1	4541	5	000113	000113 Gallus gall
24	1394	42.0	1554	4	AF230742	AF230742 Canis fam
25	1035.8	41.2	2808	5	NVPETINP	NVPETINP
26	835	25.1	2452	6	139392	139392 Sequence
27	835	25.1	2452	6	139392	139392 Sequence
28	835	25.1	2452	6	139392	139392 Sequence
29	835	25.1	2452	6	139392	139392 Sequence
30	723.4	21.8	1815	11	006460	006460 Human retin
31	711.8	21.4	3486	5	AF102861	AF102861 cncorhyne
32	639.2	19.2	2730	5	AY098289	AY098289 Gallus 1
33	483.2	14.5	2102	10	KATKTNBLST	KATKTNBLST
34	347.2	10.4	426	9	AF044224	AF044224 Homo sapi
35	288.6	8.7	944	5	AY125075	AY125075 Datto ter
36	248	7.5	301	9	F336015S25	F336015S25
37	248	7.5	301	9	F336015S25	F336015S25
38	225.2	6.8	110188	2	AF105877	AF105877 Ratius bo
39	220.8	6.6	901	6	E12562	E12562 cDNA to mRN
40	220.8	6.6	1698	6	E12562	E12562 cDNA to mRN
41	220.8	6.6	3499	6	E12562	E12562 cDNA to mRN
42	220.8	6.6	3633	6	E12562	E12562 cDNA to mRN
43	218.8	6.6	1697	6	E09941	E09941 cDNA which s
44	207	6.2	480	9	HUMRBSY79	HUMRBSY79
45	199	6.0	320	9	F336015S16	F336015S16

# ALIGNMENTS

Result 1	HUMRAH1RA	4600 bp	mRNA	Linear	PRO 12 JUL 1995
HUMRAH1RA	Human retinoblastoma associated (RB) mRNA, complete cds.				
LOCUS					
DEFINITION	Human retinoblastoma associated (RB) mRNA, complete cds.				
ACCESSION	M33647.1				
VERSION	M33647.1				
KEYWORDS	retinoblastoma protein.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Tetrapoda; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS	Friend, S.H., Borowitz, J.M., Gerber, M.R., Wang, X.F., Redemund, E., Li, F.P., and Weinberg, R.A.				
TITLE	Deletions of a DNA sequence in retinoblastomas and mesenchymal				





```

BASE 100NT 1544 a 902 c 880 q 1523 t
ORIGIN
Query Match 99.88; Score 3315; DB 6; Length 4839;
Res: Local Similarity 100.0%, Pred. No. 0;
Matches 3315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 CGAAGAGTTCACATCTGGAGTACATTCGGAGGCTATATCAAAAGAAAAAGCAACT 68
DB 47 GCACAAATTCACATCTGGAGTACATTCGGAGGCTATATCAAAAGAAAAAGCAACT 430
QY 59 GTGGAGAACTGTATCTTTATGAGAGAGTGTAGCTAGATGTGCTTACATTTAC 128
DB 431 GTGGAGAACTGTATCTTTATGAGAGAGTGTAGCTAGATGTGCTTACATTTAC 490
QY 129 TCAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAAAGCAAT 188
DB 491 TCAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAAAGCAAT 550
QY 189 TCAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAAAGCAAT 248
DB 551 TCAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAAAGCAAT 610
QY 249 GTTTCAGATCTTCAGCAAAATGAGAGAGATGTGAGATCTATATTTGACATACAG 308
DB 611 GTTTCAGATCTTCAGCAAAATGAGAGAGATGTGAGATCTATATTTGACATACAG 670
QY 309 TACGAGTACATCTACCAAAATGAGAGAGATGTGAGATCTATATTTGACATACAG 368
DB 671 TACGAGTACATCTACCAAAATGAGAGAGATGTGAGATCTATATTTGACATACAG 730
QY 369 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 428
DB 731 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 790
QY 429 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 488
DB 791 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 850
QY 489 ATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 548
DB 851 ATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 910
QY 549 TACGAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 608
DB 911 TACGAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 970
QY 609 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 668
DB 971 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 1030
QY 669 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 728
DB 1031 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 1090
QY 729 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 788
DB 1091 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 1150
QY 789 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 848
DB 1151 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 1210
QY 849 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 908
DB 1211 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 1270
QY 909 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 968
DB 1271 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 1330
QY 969 TATATACAGTACAGAAAAACATACAAATCAGTCTCCATTAATCTTTTAACTTAA 1028

```









QY 2169 GATTCGAGGCAACATATATATTCACCCGGAACAGACGACATATAAAAATTCAGAAAG 2228  
|||||  
DB 2541 GATTCGAGGCAACATATATATTCACCCGGAACAGACGACATATAAAAATTCAGAAAG 2590  
|||||  
QY 2229 TCTGATACACACCAATAMAAATGATCGCAATATAMAAATGATCGCAATATAMAAATGATCGCAAT 2288  
|||||  
DB 2591 TCTGATACACACCAATAMAAATGATCGCAATATAMAAATGATCGCAATATAMAAATGATCGCAAT 2650  
|||||  
QY 2289 ATTGAGGCAATCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2348  
|||||  
DB 2651 ATTGAGGCAATCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2710  
|||||  
QY 2349 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2408  
|||||  
DB 2711 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2770  
|||||  
QY 2409 TATTCAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGAT 2468  
|||||  
DB 2771 TATTCAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGAT 2830  
|||||  
QY 2459 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2528  
|||||  
DB 2831 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2890  
|||||  
QY 2529 TATTCAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGAT 2588  
|||||  
DB 2891 TATTCAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGAT 2950  
|||||  
QY 2589 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2648  
|||||  
DB 2951 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3010  
|||||  
QY 2649 TATTCAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGAT 2708  
|||||  
DB 3011 TATTCAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGATCATATGAGAGAT 3070  
|||||  
QY 2709 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2768  
|||||  
DB 3071 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3130  
|||||  
QY 2769 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2828  
|||||  
DB 3131 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3190  
|||||  
QY 2829 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2888  
|||||  
DB 3191 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3250  
|||||  
QY 2889 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 2948  
|||||  
DB 3251 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3310  
|||||  
QY 2949 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3008  
|||||  
DB 3311 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3370  
|||||  
QY 3009 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3068  
|||||  
DB 3371 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3430  
|||||  
QY 3069 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3128  
|||||  
DB 3431 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3490  
|||||  
QY 3129 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3188  
|||||  
DB 3491 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3550  
|||||  
QY 3189 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3248  
|||||  
DB 3551 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3610  
|||||  
QY 3249 GTTCAGAAAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAG 3308  
|||||

DB 3611 ACCTCAGAAATGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 3670  
|||||  
QY 3309 AGGATTATGATAGT 3323  
|||||  
DB 3671 AGGATTATGATAGT 3685  
|||||  
RESULT 4  
105311  
LOCUS 105311  
DEFINITION Sequence 1 from Patient EP 0259031.  
ACCESSION 105311  
VERSION 105311.1 GI:591083  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4597)  
AUTHORS Dey, J. L. P. and Friend, S.  
TITLE Human DNA in the diagnosis of retinoblastoma  
JOURNAL Patient EP 0259031-A2 1 09-MAR-1988;  
FEATURES  
Location/Qualifiers  
1..4597  
BASE COUNT 1489 a 842 c 812 g 1454 t  
ORIGIN  
Query Match 99.7%, Score 3313.4, DB 6, Length 4597;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 3314; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 9 GAGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 68  
|||||  
DB 244 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 293  
|||||  
QY 9 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 128  
|||||  
DB 294 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 353  
|||||  
QY 129 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 188  
|||||  
DB 354 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 413  
|||||  
QY 189 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 248  
|||||  
DB 414 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 473  
|||||  
QY 249 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 308  
|||||  
DB 474 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 533  
|||||  
QY 309 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 368  
|||||  
DB 534 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 593  
|||||  
QY 369 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 428  
|||||  
DB 594 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 653  
|||||  
QY 429 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 488  
|||||  
DB 654 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 713  
|||||  
QY 489 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 548  
|||||  
DB 714 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 773  
|||||  
QY 549 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 608  
|||||  
DB 774 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 833  
|||||  
QY 609 GAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGTCTGAGAGAGT 668  
|||||



















DB 182 TGTGAACTTATATATTCACACAAACAGGACCTTCGATATCTACTCAAAATAAATCTTGCA 246  
QY 140 TTGGTGGTAAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 399  
DB 247 TTGGTGGTAAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 306  
QY 400 CAAGATGATCTGGGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 459  
DB 487 GAAATGATCTGGGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 366  
QY 140 CTCTGATCTGGGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 519  
DB 467 CTCTGATCTGGGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 426  
QY 520 TCACTCTGAACATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 579  
DB 427 TCACTCTGAACATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 486  
QY 580 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 639  
DB 487 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 546  
QY 140 AAAAAATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 699  
DB 547 AAAAAATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 606  
QY 700 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 759  
DB 467 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 666  
QY 760 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 819  
DB 467 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 726  
QY 820 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 879  
DB 727 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 786  
QY 880 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 939  
DB 887 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 846  
QY 940 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 999  
DB 947 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 906  
QY 1000 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1059  
DB 907 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 966  
QY 1060 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1119  
DB 1067 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1026  
QY 1120 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1179  
DB 1127 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1086  
QY 1180 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1239  
DB 1187 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1146  
QY 1240 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1299  
DB 1247 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1206  
QY 1300 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1359  
DB 1307 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1266  
QY 1360 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1419  
DB 1367 AATGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1326

DB 1267 TACAAAGATCTGAAGATTTCTTGATATACATTTTATTAGCTAAAGAGGAGTATTACAAATG 1326  
QY 1420 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1479  
DB 1327 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1386  
QY 1480 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1539  
DB 1387 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1446  
QY 1540 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1599  
DB 1447 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1506  
QY 1600 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1659  
DB 1507 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1566  
QY 1660 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1719  
DB 1567 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1626  
QY 1720 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1779  
DB 1627 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1686  
QY 1780 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1839  
DB 1687 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1746  
QY 1840 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1899  
DB 1747 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1806  
QY 1900 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1959  
DB 1807 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1866  
QY 1960 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2019  
DB 1867 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1926  
QY 2020 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2079  
DB 1927 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1986  
QY 2080 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2139  
DB 1987 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2046  
QY 2140 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2199  
DB 2047 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2106  
QY 2200 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2259  
DB 2107 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2166  
QY 2260 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2319  
DB 2167 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2226  
QY 2320 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2379  
DB 2227 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2286  
QY 2380 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2439  
DB 2287 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2446  
QY 2440 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2499  
DB 2347 TATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2406





[illegible]











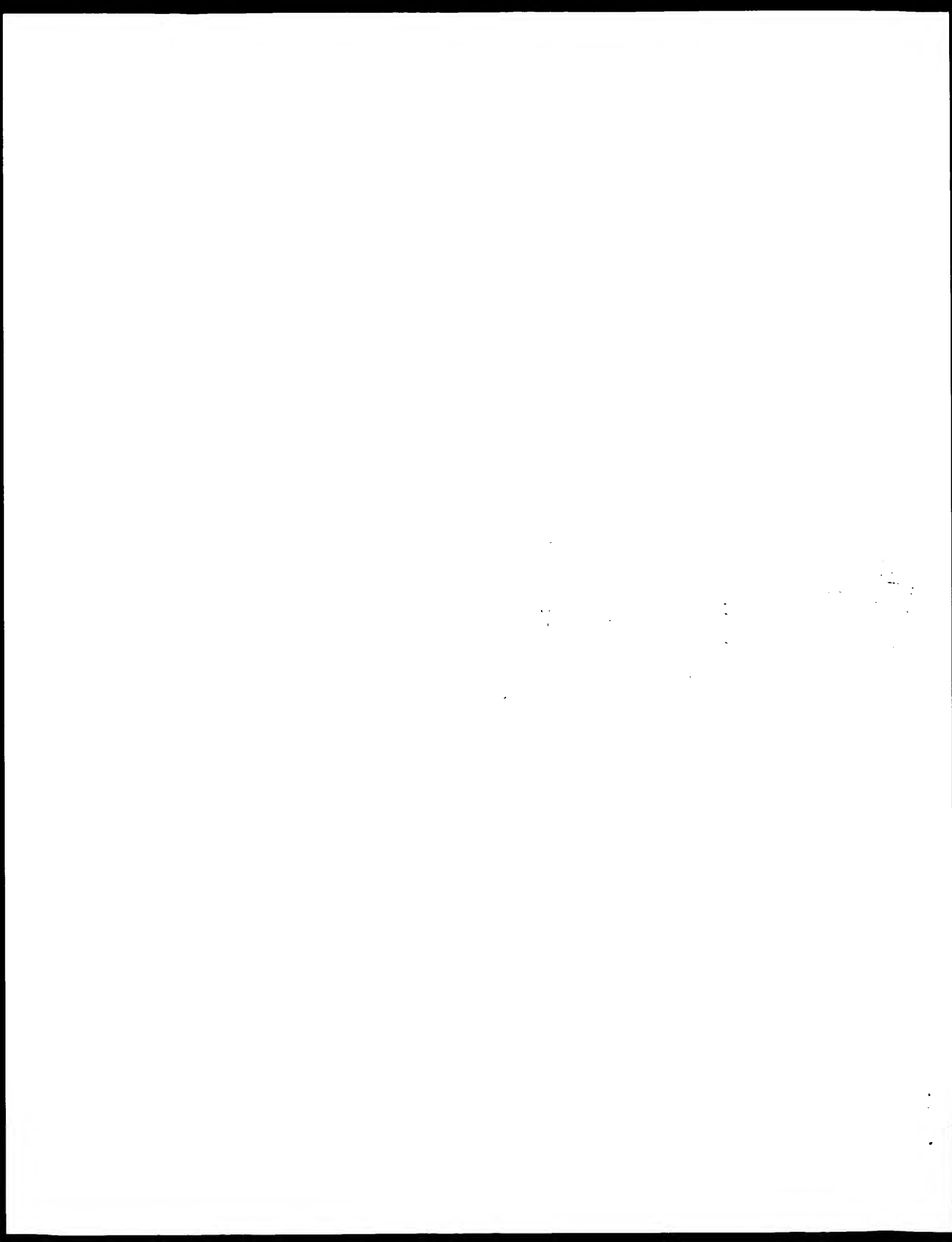












Genforce version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein nucleic search, using frame\_plus\_pzn model

Run on: January 17, 2003, 18:15:26 : Search time 4608.78 Seconds  
(without alignments)  
6678.212 Million cell updates/sec

Files: us-09-026-459a-31  
Perfect score: 4504  
Sequence: 1 MDTALQKLIKPIHVERA.....TRMKKMNUSMITSNKEEK 874

Scoring table: BLASTN62  
Xgapop 10.0 : Xgapext 0.5  
Ygapop 10.0 : Ygapext 0.5  
Fgapop 6.0 : Fgapext 7.0  
Delop 6.0 : Delext 7.0

Searched: 2054540 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4106280

Minimum DB seq length: 0

Maximum DB seq length: 60000000

Post processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

Model frame: pzn model -DEXT-ylh  
-pzn 2.0 -pzn 2.0 -pzn 2.0 -pzn 2.0 -pzn 2.0 -pzn 2.0 -pzn 2.0 -pzn 2.0 -pzn 2.0 -pzn 2.0  
Units: bits -START=1 -END=1 -MATRIX=BLASTN62 -TRANS=human40.cdi -LIST=45  
LOCALIGN=200 -THR\_SCORE=100 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
OUTFILE=pro -NORM=ext -HAPSI=500 -MINLEN=0 -MAXLEN=20000000  
NO\_XLIFY NO\_XLIFY -LAP=QUERY -NEG\_SCORES=0 -WAP=1000000 -DEXT=120  
WARN TIMEOUT 40 THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -F5GAPEXT=7  
YGAPOP=10 YGAPEXT=0.5 DELOP=6 DELEXT=7

Database:

- 1: db\_bact\*
- 2: db\_btq\*
- 3: db\_btq\*
- 4: db\_btq\*
- 5: db\_btq\*
- 6: db\_btq\*
- 7: db\_btq\*
- 8: db\_btq\*
- 9: db\_btq\*
- 10: db\_btq\*
- 11: db\_btq\*
- 12: db\_btq\*
- 13: db\_btq\*
- 14: db\_btq\*
- 15: db\_btq\*
- 16: db\_btq\*
- 17: db\_btq\*
- 18: db\_btq\*
- 19: db\_btq\*
- 20: db\_btq\*
- 21: db\_btq\*
- 22: db\_btq\*
- 23: db\_btq\*
- 24: db\_btq\*
- 25: db\_btq\*
- 26: db\_btq\*
- 27: db\_btq\*
- 28: db\_btq\*

- 29: em\_btq\_hum\*
- 30: em\_btq\_hum\*
- 31: em\_btq\_hum\*
- 32: em\_btq\_hum\*
- 33: em\_btq\_hum\*
- 34: em\_btq\_hum\*
- 35: em\_btq\_hum\*
- 36: em\_btq\_hum\*
- 37: em\_btq\_hum\*
- 38: em\_btq\_hum\*
- 39: em\_btq\_hum\*
- 40: em\_btq\_hum\*
- 41: em\_btq\_hum\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4499	99.9	2994	6	AKG91965 Sequence
2	4499	99.9	2994	6	AKG91965 Sequence
3	4499	99.9	2994	6	AKG91965 Sequence
4	4499	99.9	2994	6	AKG91965 Sequence
5	4499	99.9	2994	6	AKG91965 Sequence
6	4499	99.9	2994	6	AKG91965 Sequence
7	4499	99.9	2994	6	AKG91965 Sequence
8	4499	99.9	2994	6	AKG91965 Sequence
9	4499	99.9	2994	6	AKG91965 Sequence
10	4499	99.9	2994	6	AKG91965 Sequence
11	4499	99.9	2994	6	AKG91965 Sequence
12	4499	99.9	2994	6	AKG91965 Sequence
13	4499	99.9	2994	6	AKG91965 Sequence
14	4499	99.9	2994	6	AKG91965 Sequence
15	4499	99.9	2994	6	AKG91965 Sequence
16	4499	99.9	2994	6	AKG91965 Sequence
17	4499	99.9	2994	6	AKG91965 Sequence
18	4499	99.9	2994	6	AKG91965 Sequence
19	4499	99.9	2994	6	AKG91965 Sequence
20	4499	99.9	2994	6	AKG91965 Sequence
21	4499	99.9	2994	6	AKG91965 Sequence
22	4499	99.9	2994	6	AKG91965 Sequence
23	4499	99.9	2994	6	AKG91965 Sequence
24	4499	99.9	2994	6	AKG91965 Sequence
25	4499	99.9	2994	6	AKG91965 Sequence
26	4499	99.9	2994	6	AKG91965 Sequence
27	4499	99.9	2994	6	AKG91965 Sequence
28	4499	99.9	2994	6	AKG91965 Sequence
29	4499	99.9	2994	6	AKG91965 Sequence
30	4499	99.9	2994	6	AKG91965 Sequence
31	4499	99.9	2994	6	AKG91965 Sequence
32	4499	99.9	2994	6	AKG91965 Sequence
33	4499	99.9	2994	6	AKG91965 Sequence
34	4499	99.9	2994	6	AKG91965 Sequence
35	4499	99.9	2994	6	AKG91965 Sequence
36	4499	99.9	2994	6	AKG91965 Sequence
37	4499	99.9	2994	6	AKG91965 Sequence
38	4499	99.9	2994	6	AKG91965 Sequence
39	4499	99.9	2994	6	AKG91965 Sequence
40	4499	99.9	2994	6	AKG91965 Sequence
41	4499	99.9	2994	6	AKG91965 Sequence
42	4499	99.9	2994	6	AKG91965 Sequence
43	4499	99.9	2994	6	AKG91965 Sequence
44	4499	99.9	2994	6	AKG91965 Sequence
45	4499	99.9	2994	6	AKG91965 Sequence

ALIGNMENTS

RESULT 1











```

CY 722 ProThrLeuSerProIleProHisIleProArqSerProTyrLysPheProSerSer 741
DB 2444 CCCCCIACCTIGACCAATACCHACATCCCTCGAACCCCTTACAAAGTTCCCTACTTCA 2523
CY 742 ProLeuArqIleProGlyGlyAsnIleTyrLysSerProLeuLysSerProTyrLysIle 761
DB 2624 CCCCCIACCTIGACCAATACCHACATCCCTCGAACCCCTTACAAAGTTCCCTACTTCA 2683
CY 762 SerGluGlyLeuProThrProThrLysMetThrProArqSerArqIleLeuValSerIle 781
DB 2904 TCAGAGAGCTGCGCAACACCAACCAAAATGACTCCCAAGATCAAGAATCTTAGTATCAAT 2643
CY 782 GlyIleSerPheGlyThrSerGlyLysPheGlyLysIleAsnGlyMetValCysAsnSer 801
DB 2644 CCCCCIACCTIGACCAATACCHACATCCCTCGAACCCCTTACAAAGTTCCCTACTTCA 2703
CY 802 AspArqValLeuLysArqSerAlaGlySerAsnProProLysProLeuLysLysLeu 821
DB 2704 CACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2763
CY 822 ArgGlyAspIleGlyGlySerAspGlyLysIleAspGlySerLysHisLeuProGlyGlySer 841
DB 2904 CCCCCIACCTIGACCAATACCHACATCCCTCGAACCCCTTACAAAGTTCCCTACTTCA 2923
CY 842 LysPheGlnGlnLysLeuAlaGlyMetThrSerThrArqMetThrArqMetGlnLysGlnLys 861
DB 2824 AAAATTCACCAACCTGACCAAAATGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2883
CY 862 MetAsnAspSerMetAspThrSerAspLysGluGlyLys 874
DB 2804 ATGAATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2922

RESIDUE 4
LENGTH 28
DEFINITION Tissue specific expression of 1-4 in human prostate
ACCESSION BD009728
VERSION BD009728.1 GI:18638101
KEYWORDS JP 2001504638-A/2.
SOURCE unidentifed.
ORGANISM unidentifed.
REFERENCE 1 (bases 1 to 2994)
AUTHORS Antomadeo, Gregory, P. I. and Willis, K. N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL CANCER 100:110-115 1997
COMMENT OS unidentifed
PN JP 2001504638-A/2
PD 21-MAR-2001
PR 15-NOV-1997 JP 1998522958
DEPOSITED 15-NOV-1997 IN 09/2001/002
DEPOSITED IN ANTIMAN, PICHARD, J. GREGORY, KENNETH N. WILLIS, P. C.
CC Strandedness: Single;
CC Topology: Linear;
FH KEY location/Qualifiers
FT source 1..2994
FT location/Qualifiers
FT /organism="unidentifed"/
FT /db_xref="taxon:32644"

BASE 1 (IN) 474 a 618 c 593 g 809 t
BASE 2 (IN)

Alignment Scores:
Prod. No.: 0 Length: 2994
Score: 4499.00 Matches: 873
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.89% Indels: 0

```

```

DB: 6 Gaps: 0
US-09-026-459a-31 (1-874) x HD009728 (1-2994)
CY 2 AspPheThrAlaLeuCysGlnLysLysIleProAspHisValArgGluArqAlaIleP 21
DB 304 GATTTTACTGCTATTATGTCACAAATTAAAGATATATATCATGCTGCAGAGAGAGAGTTTGA 463
CY 22 LeuPheThrProGlyLysValSerValAspGlyValLeuGlyGlyTyrIleClnLysLys 41
DB 364 TTAACCTTGGACAGAAAGTTTCATGCTGGATGGAGTATTCGGAGCTTATATTCAAACAAA 123
CY 42 LysGlnLeuIleGlyValLeuPheIleAlaAlaValAspLeuAspGlyMetSerPhe 61
DB 424 AAGCAACGTGGGCAATGCTGATCTTTATTCGACACATTCACCTAGATGAGATGCGTTC 483
CY 62 ThrPheThrGluLeuGlnLysAsnIleGluIleSerValHisLysPhePheAsnLeuLeu 81
DB 484 ACITTTACTGAGCTACASAAAAACATAGAAATCAGTGTGCTAAATTTCTTTAACTTACTA 543
CY 82 LysGluIleAspThrSerThrLysValAspAspValMetSerAlaLeuLeuLysLys 101
DB 544 AAGCAATTCATACCCAGTACCAAGTTGATATGCTATGCTCAAGACGTGTGAGAGAT 603
CY 102 AspValLeuPheAlaLeuPheSerLysLeuAlaArqThrCysGlnLeuIleTyrLeuThr 121
DB 604 CATCTATGCTTTCACCTTCACCTCAAAATGGAAAGCACAIGTCAACTTATATATGACA 663
CY 122 ClnProSerSerThrSerThrGlnIleAsnSerAlaLeuValLeuLysValSerIleP 141
DB 664 CAGCCAGAGCTTGGATATCTCTGAAATTAATTTCTGCTATGCTGCTAAAGTTTCTTGG 723
CY 142 IleThrPheLeuLeuAlaLysGlyGluValLeuGlnMetClnAspAspLeuValIleSer 161
DB 724 ATCACTTTTATTAGCTAAGGGGAGATATACAAATCGAACATGATCTGCTGATATCA 783
CY 162 PheGlnLeuMetLeuCysValLeuAspTyrPheIleLysLeuSerProProMetLeuLeu 181
DB 784 TTTCAGTTAAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843
CY 182 LysGlnProTyrLysThrAlaValIlePheIleClnAspGlySerProArqThrProArq 201
DB 844 AAGAACCATATTAACACAGCGTGTATACCCATTATTTGCTGCTGCTGCTGCTGCTGCTGCT 903
CY 202 GlyLeuAsnArqSerAlaGlyAlaLysGlyLeuGlnIleAsnAspThrArgIleClnLeu 221
DB 904 GGTTCAGAACACGAGTGCACGCTAGCAAAACAACTAGAAATCATACAGAAGATTAATCA 963
CY 222 ValLeuGlyLysGlnIleSerLysAsnIleAspGlnValLeuLysValIlePheLys 241
DB 964 GTTCTGCTGAAACAAACATCAATGCTAATATAGATGAGGCTGCAAAATGCTTATTCAAA 1023
CY 242 PheIleThrPheMetAspSerLeuGlyLeuValThrSerAspGlyLeuThrGluVal 261
DB 1024 TTTAATCTTTATTAATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
CY 262 AsnLeuSerLysArqTyrGluGlnIleTyrLeuLysAsnLysAspLeuAspAlaArqLeu 281
DB 1084 AATCTTTCTAAACGATACGAAGAAATTTATCTTAAAAATAAAGATCTAGATCAACA 1143
CY 282 PheLeuAspHisAspLysThrLeuGlnIleThrAspSerIleAspSerPheGluThrGln 401
DB 1144 TTTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
CY 302 ThrProArgLysSerAsnLeuAspGlyGluValAspValIleProGlnHisThrProVal 421
DB 1204 ACACCAAGAAAAAGTAACTTGAATGAAGAGGAGGATGATGATGATGATGATGATGATGAT 1263
CY 322 ArgThrValMetAsnThrIleGluGlnLeuMetMetIleLeuAsnSerAlaSerAspGln 341
DB 1264 AGGACGTGTAAGAACATATCCCAACAAATTAATGATGATTTAAATTCACCAAGATCA 1323
CY 342 ProSerGluAsnLeuIleSerTyrPheAspAsnGlyThrValAsnProLysGluSerIle 361

```







CY	602	ThrProValIysSerAsnLeuAspGluValAsnValHisProProHisHisProVal	321
DB	1284	AACCAAGAAAGAATCTTGTATGAAGAGTGAAATGGTAATTCCTGCACACCCAGTT	1263
CY	622	ArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGln	341
DB	1264	AGACAGCATTACGACACTATCCAACAAATAATGATGATTATAATTCGACCAAGCATCAA	1323
CY	642	ProSerGluAsnLeuIleSerTyrrPheAsnAsnCysThrValAsnProLysGluSerIle	361
DB	1324	CCCTTCAGAAAATCGAATTCCTATTTTAACAACTCCACAGTCATCCCAAGAAAGATATA	1383
CY	662	IsofYSAWVallySASEIleGlyTyrIleChelYsGluLysPheAlaLysAlaValIGLY	381
DB	1384	CTCAAAAACGTAAACATATACATACATCTTTAAAGAGNAATTTCTAAAGCTGTGGGA	1443
CY	682	GlnGlyCysValGluIleGlySerGlnArgTyrLysLeuGlyValArgLeuGlyTyrTyrArg	401
DB	1444	CAGGGTGTGTGCCAAATGGCATCAACAGCAACAACATTCGAGTTCCTCTCTATACCGA	1503
CY	702	ValMetGluSerMetLeuLysSerIleGluArgLeuSerIleGlnAsnPheSerLys	421
DB	1504	GTAAATGCAATCATCTCTTAATTCAGAAACAGACCAATATGCCATTCAAAAATTAACAAA	1563
CY	722	IsoLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValIleMet	441
DB	1564	CTCTCAATGCAACAAATTTTCATATGCTTTATGGGTGGCTCTTAGGTTGTAAATG	1623
CY	742	AlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro	461
DB	1624	GTCAATATAGCAGAGATACATCTCAGAACTCTGATCTCGAACACAGATTGTCTTTCCCA	1683
CY	762	TyrIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPhe	481
DB	1684	TCGATTCGCAATCTCTTAATTTAAAGCCCTTCATTTTTACAAAGTCATCGAAAGTTT	1743
CY	782	IleLysAlaGluGlyAsnLeuThrArgGluMetIleLysHisLeuGluArgCysGluHis	501
DB	1744	ATCAAAACAAAGAGANATTGAAACAAATATATAAANAATTTAAAGATGTGAACAT	1803
CY	802	ArgIleMetGluSerLeuAlaIleTrpLeuSerAspSerProLeuPheAspLeuIleLysGln	521
DB	1804	GSAAATCAAGATTCCTTCGATGGTCTCCAGATTCACCCTTATTTCATCTATAAACAA	1863
CY	822	SerLysAspArgLeuLeuProThrAspHisLeuIleLeuSerAlaCysProLeuAsnLeuPro	541
DB	1864	TAAAGACATCACAAAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	1923
CY	842	IsofYSAWVallySASEIleGlyTyrIleChelYsGluLysPheAlaLysAlaValIGLY	561
DB	1924	CTCCAGAAATATCACATGTCACACATATGATCTTCTCTGTGTAGATCTCCAAAGMA	1983
CY	862	LysGlySerIleIleArgValAsnSerIleAlaAsnAlaGluThrGlnAlaIleSerAla	581
DB	1984	AAAGSTGTCACTACGCGGTAAATTCATCTGCAATCAATCAATCAATCAATCAATCAAT	2043
CY	882	PheGlnTrpGlyLysProLeuIleYsGlnThrSerLeuSerIlePheTyrIleValTyr	601
DB	2044	TCCACAGCCCAAGACCCATTCAAAATCCTCTCTTTCACATGTTTTATAAAAAGTGTAT	2103
CY	902	ArgLeuAlaIleTyrLeuArgLeuAsnThrLeuCysGlnArgIleLeuSerGlnHisProGln	621
DB	2104	CGAGTAAAGTATCTCGGCTAAATACATTTGTGAAGCGCTTGTGTGACACACAGAA	2163
CY	922	IsofYSAWVallySASEIleGlyTyrIleChelYsGluLysPheAlaLysAlaValIGLY	641
DB	2164	TTACAAAT	2223
CY	942	ArgAspArgHisLeuAspGlnIleMetLysSerMetTyrGlyIleCysLysValLys	661
DB	2224	ACAGACAGCAATTCGACCAAAATACAGAGCTTCCAATGATGATGATGATGATGATGAT	2283



2 AspPheThrAlaLeuGlySerLeuValLysPheProAspHisValAlaArgAlaIleTrp 21  
 106 GATTTATTCATATGATGCAAAATAAAGATACATCATCATGTCACACAGACGCTGG 225  
 22 LeuThrTrpGlyLysValSerGlyValLeuGlyGlyTrpPheLeuLysLys 41  
 226 TTAATTCGGGAAVAAAGHTTCATGCTGGGAGGATATGGGAGGTATATTCAAAAGAA 285  
 42 LysLeuLeuTrpGlyLysCysLeuPheLeuAlaValAspLeuMetSerPhe 61  
 286 AAGAAATGAGGAAATGTGATCTTTATTCAGCAATGCACTAGAGATGGCTTC 145  
 62 ThrThrThrGluGluGluLysAsnLeuLeuIleSerValHisLysPheAsnLeuLeu 81  
 346 AATTTTACGATACACAAAATAACAAAACACAGCTGCATCAAAATTTTAACTTAC 405  
 82 LysLeuIleLeuAspThrSerThrLysValAspAsnAlaMetSerAlaLeuLeuLysLysTrp 101  
 406 AAAAAATGATATCAATGAAATGATATGATGATGATGATGATGATGATGATGAT 465  
 102 AspValLeuPheAlaLeuPheSerLysLeuGlyArgThrCysGluLeuIleTrpLeuThr 121  
 466 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525  
 122 GluProSerSerIleSerThrGluIleAsnSerAlaLeuValLeuLysValSerTrp 141  
 526 AAAAAATGATATCAATGAAATGATATGATGATGATGATGATGATGATGATGAT 585  
 142 ThrThrPheLeuAlaLysGlyValLeuGluMetGluAspAspLeuValIleSer 161  
 586 AATCATTTTATATAGTAAAGGAGAGATATCAAAATGAGATGATGATGATGATGAT 645  
 162 PheGluLeuMetLeuGlyValLeuAspTrpPheIleLysLeuSerProMetLeuLeu 181  
 646 TTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 705  
 182 LysLeuProTrpLysThrAlaValIleProIleAsnGlySerProArgThrProArgArg 201  
 706 AAAAAATGATATCAATGAAATGATATGATGATGATGATGATGATGATGATGAT 765  
 202 GlyValAsnArgSerAlaArgIleAlaLysGluMetGluAspThrArgIleLeuGlu 221  
 766 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 825  
 222 ValLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 241  
 826 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 885  
 242 PheIleProProMetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGlu 261  
 886 TTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 945  
 262 AsnLeuSerLysArgTrpGlyLeuIleLysLysLysLysLysLysLysLysLys 281  
 946 AATTTTCAAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1005  
 282 PheLeuAspHisLysLysThrLeuGluThrAspSerIleAspSerPheGluThrGluArg 301  
 1006 TTTTGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065  
 302 ThrProArgLysSerAsnLeuAspLeuValAlaValIleProProHisThrProVal 321  
 1066 AAAAAATGATATCAATGAAATGATATGATGATGATGATGATGATGATGATGAT 1125  
 322 ArgThrValMetAsnThrIleGluGluIleMetMetIleLeuAsnSerAlaSerAspGlu 341  
 1126 AAAAAATGATATCAATGAAATGATATGATGATGATGATGATGATGATGATGAT 1185  
 342 ProSerGluAsnLeuIleSerTrpPheAsnAsnCysThrValAsnProLysGluSerIle 361  
 1186 GTTCAAAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1245

962 LeuLysArgValLysAspIleLeuTrpTrpIlePheGlyGlyIleAlaValValVal 381  
 1246 CTGAAACAGTCAAGGATATAGGATACATCTTAAACAGAGAAATTTCTAAACATGGTGGCA 1305  
 382 GluGlyCysValGluIleGlySerGluArgTrpLysLysLysLysLysLysLysLys 401  
 1306 CAGGTTTGTTGGTGAATTTGGATCAATAGGATATAAATTTGAGTTTGTCTGATATATGGA 1365  
 402 ValMetGluSerMetLeuLysSerGluGluGluArgLeuSerIleGluAsnProSerLys 421  
 1366 GTAAATGGAATCAATGCTTAAATCAGAGAGAGAGAAATTAATCAATTTAAATTTAGCAAA 1425  
 422 LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaLysAlaLeuValValMet 441  
 1426 CTCTCTGAAATGCAACATTTTCAATATGCTTATTTGAGGAGGCTTCTGAGGTTGAAATG 1485  
 442 AlaThrTrpSerArgSerThrSerGluAsnLeuAspSerGlyPheAspLeuSerPhePro 461  
 1486 GGCATATATAGCAAGATATATCTCAGATGATCTTCAATTTGAGAGAAATTTGCTTTTTCGA 1545  
 462 TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTrpLysValIleLeuLeuSerPhe 481  
 1546 TCGATTCTGCAATGCTTAAATTTAAAGCTTTTCAATTTTACAAAGCTGATGAGAACTTT 1605  
 482 IleLysAlaGluGlyAsnLeuThrArgGluMetIleLysHisLeuGluAlaGlyGluHis 501  
 1606 ATCAAGCAGAGAGCAATTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1665  
 502 ArgIleMetGluSerLeuAlaTrpLeuSerAspSerThrLeuPheAspLeuIleGlyHis 521  
 1666 GATCATGGAATGCTTGTGATGCTTGTGATGCTTGTGATGCTTGTGATGCTTGTGATGCTT 1725  
 522 SerLysAspArgGluGlyProThrAspHisLeuGluSerAlaLysProLeuAsnLeuPro 541  
 1726 TCAAGGAGCAG 1785  
 542 LeuGluAsnAsnHisThrAlaAlaAspMetTrpLeuSerProValArgSerProLysLys 561  
 1786 CTCCAGAAAT 1845  
 562 LysGlySerThrThrArgValAsnSerThrAlaAsnAlaThrThrGluAlaThrSerAla 581  
 1846 AAGATTCATATAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1905  
 582 PheGluThrGluLysProLeuLysSerThrSerLeuSerLeuPheTrpLysLysValTrp 601  
 1906 TTCCAGAGCCAG 1965  
 602 ArgLeuAlaTrpLeuArgLeuAsnThrLeuCysGluArgLeuLeuSerGluHisProGlu 621  
 1966 CGGTAGGCTTATTCGGGCTAAATATACCTTGGGAAATGCTTGTGATGATGATGATGATGAT 2025  
 622 LeuGluHisIleIleTrpThrLeuPheGluHisThrLeuGluAsnGluTrpLeuLeuMet 641  
 2026 TTAAACAT 2085  
 642 ArgAspArgHisLeuAspGluIleMetMetCysSerMetTrpGlyIleLysLysValLys 661  
 2086 ACACAGAGCATTTGGAGCAAAATATATATATATATATATATATATATATATATATATAT 2145  
 662 AsnIleAspLeuLysPheLysIleIleValThrAlaTrpLysAspLeuProHisAlaVal 681  
 2146 AATATAGAGCTTAAATTTCAAAATATATATATATATATATATATATATATATATATAT 2205  
 682 GluGluThrPheLysArgValLeuIleLysGluGluGluTrpAspSerIleIleValPhe 701  
 2206 CAGAGAGATTCAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2265  
 702 TyrAsnSerValProMetGluArgLeuLysThrAsnIleLeuGluGluTrpAlaSerThrArg 721  
 2266 TATAACATGCTCTTCAATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2325  
 722 ProProThrLeuSerProIleProHisIleProArgSerProTrpLysPheProSerPro 741





FEATURES  
source location/offsets  
1..4839

/organism: "Homo sapiens"  
/db\_xref: "taxon:9606"

RMS: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1

query: 15.4 3 302 880 4 153 1













```

1144 TTTTGGATCATGATTAATAAATCTTTAGACTGATGTTTATAGACAGAGTTTGAAGACACAGAGA 1203
QY 602 ThrProAlaGlySerAsnLeuAspGluValAsnValIleProHisThrProVal 321
1204 ACAATATGAAAGTAACCTTGAAGACAGCTGAATTAATCTTCCACACACATCCAGTT 1263
QY 622 ArgThrValMetAsnThrIleGluGluLeuMetIleLeuAsnSerAlaSerAspGln 341
1264 AGGACGCTTATCAACACATATCCACAAATTAATCATGATTTTAAATCCACCAAGGATCAA 1323
QY 642 ProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProIysGluSerIle 361
1324 GCTTCAGAAAATCTGATTTGCTATTTTAACTGACAGCTGATGATGCAAAAGAAATATA 1383
QY 662 LeuIysAsnValGlyAspIleCysThrIleHisLeuGlySerPheAlaValGly 381
1384 CAGAAAAGAGTCAAGAAATATAGGATATATCTTAAAGACAGAAATTCGAAAGGCTGGCA 1443
QY 682 GlyGlyValGlyIleGlySerGluArgTyrIleLeuGlyValAlaGlyLeuTyrArg 401
1444 CAGGGTCTGTCGAAATTCGATCAGCGCATACAAACTTCGAGTTCCGTGTATTAACCA 1503
QY 692 ValMetGluSerMetLeuIysSerGluGluGluArgLeuSerIleGluAsnPheSerIys 421
1504 GATATGGAATCCATCTTAAATCAGAGAGAGAGGATTTCCATTCCTCAAAATTTTACAAA 1563
QY 712 LeuLeuAsnAspAsnIlePheHisMetSerIleLeuAlaCysAlaLeuGluValAlaMet 441
1564 CTCTTGAATGACAACTTTTCAATATGCTTTATTCGGGCTGGCTTCGAGGTGTGATG 1623
QY 732 AlaThrTyrSerArgSerThrSerGluAsnLeuAspSerGlyThrAspLeuSerPhePro 461
1624 GGCATATATACAGCAAGTACATCTCAGAAATCTTCATTTCTGCAACAGATTTGCTTTCCCA 1683
QY 752 ThrIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerPhe 481
1684 TGGATTTGCAATGCTTAAATTTAAGACCTTTCATTTTACAAATGATGAGAACTTTT 1743
QY 772 IleIysAlaGluGlyAsnLeuIleArgGluMetIleIysHisLeuGluArgCysLeuHis 501
1744 ATCAAAATATAGAAATCAATCTTCAATATGAAATGATATGATATGATATGATATGATAT 1803
QY 792 ArgIleMetGluSerLeuAlaTyrLeuSerAspSerProLeuPheAspLeuIleIysGln 521
1804 GCAATATGGAATGCTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1863
QY 812 SerIysAspArgGlyGlyProThrAspHisLeuSerAlaCysProLeuAsnLeuPro 541
1864 TCAAAAGACCGAAGACGAAAGAAATGATATGCTTGAATCTGCTGCTGCTGCTGCTGCTG 1923
QY 832 LeuGluAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProIysIys 561
1924 CTCACAAATATATCATATGACATATGATATGATATGATATGATATGATATGATATGAT 1983
QY 852 TysIysSerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAla 581
1984 AAAATTTCAACACGCTGATAAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2043
QY 872 PheGluThrGlnIysProIysGlySerThrSerLeuSerLeuPheTyrIysIysValTyr 601
2044 TTGAGAGACAAAGACCAATGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103
QY 892 ArgIleAlaIleTyrLeuArgLeuAsnThrIleCysGluArgLeuLeuSerGluHisProGlu 621
2104 GGGATAGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2163
QY 912 LeuAlaHisIleIleThrThrIlePheGluHisThrLeuGlnAsnLeuTyrGluLeuMet 641
2164 TTAACAATAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2223
QY 932 ArgAspArgHisLeuAspGlnIleMetMetCysSerMetTyrGlyIleCysIysValIys 661
2224 AGACAAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2283

```

```

QY 662 AsnIleAspLeuLysPheLysIleIleValIleAlaTyrIleLysAspLeuProHisAlaVal 681
1284 AATATACACCTTAAATTTCAAAATCAATTTGTAACAGCA ACAAGGATCTCTCTCAIGCTG 2343
QY 682 GlnGluThrPheLysArgValLeuIleLysGluGluIleTyrAspSerIleIleValPhe 701
2344 CAGAGACATTCAGAAAGTGTTCATCAAAAGAGAGAGATATGATCTTATTATATATTC 2403
QY 702 TyrAsnSerValPheMetGluArgLeuLysIleAsnIleLeuGlnIleAlaSerThrAla 721
2404 TATATCTGATCTCTCATGAGAGACTGAAACAAATATTTGAGATGATGTTTCAGAGAG 2463
QY 722 ProProThrLeuSerPheIleProHisIlePheArgSerPheTyrLysPheProSerSer 741
2464 CCCCCTACTTGTGACCAAACTTCATCAATTCCTGAGAGAGAGAGAGAGAGAGAGAGAG 2523
QY 742 ProLeuArgIleProGlyGlyAsnIleTyrIleSerPheLeuLysSerPheTyrIleVal 761
2524 CCGTTATGATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2583
QY 762 SerGluGlyLeuArgThrIleThrIleLysMetIlePheArgSerArgIleLeuValSerIle 781
2584 TCAAGATTTTGTCAAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2643
QY 782 GlyGluSerPheGlyThrSerGluLysPheGlyIleAsnGlnMetValCysAsnSer 801
2644 CCTCAATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2703
QY 802 AspArgValIleLysArgSerAlaGlyLysSerAsnGlyProGlySerLeuLysLysLeu 821
2704 GAGCGTGTCTTAAATAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2763
QY 822 ArgPheAspIleGlyGlySerArgAlaAspCysLeuHisIleSerIleVal 841
2764 CCGTTTATATTTAGAGATAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2823
QY 842 LysPheThrIleLysLeuAlaGlyMetThrSerThrArgThrArgMetGlnLysIysLeu 861
2824 AAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2883
QY 862 MetAsnAspSerMetAspThrSerAsnLysGluLys 874
2884 ATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2922

```

## RESULT 13

## HUMBERS

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITF

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

## COMMENT

HUMBERS 4746 bp mRNA linear PRI 12-JUL-1995  
Human retinoblastoma susceptibility mRNA, complete cds.

DEFINITION Human retinoblastoma susceptibility mRNA, complete cds.  
ACCESSION M15400  
VERSION M15400.1 GI:190958

KEYWORDS retinoblastoma susceptibility.  
SOURCE Homo sapiens (clone: RB [1.5]) foetus retina cDNA to mRNA.

ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 243 to 4740)  
Lee, W.H., Bookstein, R., Hong, P., Young, L.J., Shew, J.Y., and Lee, J.Y.  
Human retinoblastoma susceptibility gene: cloning, identification,  
and sequence

JOURNAL Science 235 (4794), 1394-1399 (1987)  
MEDLINE 87149066  
PUBMED 3823889

REFERENCE 2 (bases 1 to 480)  
Lee, W.H., Bookstein, R., Young, L.J., Lin, C.J., Rosenfeld, M.J., and  
Lee, W.H.

TITF Molecular mechanism of retinoblastoma gene inactivation in  
retinoblastoma cell line Y79

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (16), 6017-6021 (1988)  
MEDLINE 88320373  
PUBMED 3413073

COMMENT Draft entry and computer-readable copy of sequence in [1] kindly  
provided by F Bookstein, 27 APR-1987.









27	655	GluAspGluTyrGluGluMet	AtLASP1	ABISL	AspGluGluMet	AtLASP1	654
46	1747	CAGAAATCACTAAGAACTCA	AtLASP1	AGCA	AGCACT	AtLASP1	1806
27	656	TyrGlyTyrGlySerValTyrAsp	AtLASP1	ValTyrAsp	ValTyrAsp	AtLASP1	674
46	1807	TATGGATATGCAAAAGGAA	AtLASP1	AAAG	AAAG	AtLASP1	1866
27	675	LysAspGluProHisAlaVal	AtLASP1	ValGluGluGlu	ValGluGluGlu	AtLASP1	694
46	1867	AAGATATCTGCAATGCTT	AtLASP1	CAACACAT	CAACACAT	AtLASP1	1926
27	696	TyrAspSerHisLeuValPhe	AtLASP1	ValPheMet	GluArgLeuLys	ThrAsnIle	714
46	1927	TATCATCTAATAAGATCT	AtLASP1	AACTGGCT	CTCATGACACAT	CGAAACAAATAT	1986
27	715	LeuGluTyrAlaSerThrArg	ProProHis	LeuSerPro	LeuProHis	LeuProArgSer	734
46	1987	TTGAGTATGCTCAAGAG	CTTCA	CAATATGCT	CAATATGCT	CAATATGCT	2046
27	736	ProTyrLysPheProSerPro	LeuAla	LeuGlyAsnIle	ValLeuSerPro	ValLeuSerPro	754
46	2047	CACTTCAACCTTCTAGT	CTCACT	CACTTCTAGT	CACTTCTAGT	CACTTCTAGT	2106
27	756	LeuLysSerProTyrLysHis	LeuSerGlu	GlyLeuPro	ThrProThrLysMet	ThrProArg	774
46	2107	CTGAAGATCTCAATAAAAA	CTCTG	CAACACAA	AAAAAGATCTG	CAACAA	2166
27	775	SerArgLeuValSerHisGly	LeuSerPhe	GlyThrSerGlu	LysPheGluLysHis	Leu	794
46	2167	CTAAATCTTATATCAAT	CTGCAAT	CTCTGCAAGCT	CTCTGCAAGCT	CTGCAAGCT	2226
27	796	AsnIleMetValTyrAspSer	AspAla	ValLeuLysArg	SerAlaLeuLysSerAsnPro	Arg	814
46	2227	AACTATATGATGATAAAG	CTGCTGCT	CAAAAGCT	CTGCTGCAAGCT	CTGCAAGCT	2286
27	815	ProLysProLeuLysIleAsn	PheAsp	LeuGlySerAsp	IleAlaAspLysSer	Arg	834
46	2287	CTTAAATGACGCAAAAAC	CTACCT	CTCATATGCAAGCA	CTCATGCAAGCT	CTGCAAGCT	2346
27	836	LysHisLeuProGlyGluSer	LysPheGlu	GluLysGluAla	MetThrSerThrArg	Arg	854
46	2347	AAATATCTCAAGAGATG	CAATTT	CTCAAGAACT	GCAAAATGCT	TATCTCA	2406
27	855	ThrArgMetGluLysGlnLys	MetAsn	AspSerMetAsp	ThrSerAsnLysGluLys	Arg	874
46	2407	ATCAAAATCAAAAAATCA	ATCAAT	CAATGCAATGCAAT	CAATGCAATGCAAT	CAATGCAAT	2466

Manuscript completed: January 18, 2004, 22:15:50  
 Job time: 408.01 secs











10	799	ATGTCGCTTTCACATATTTATTAATTAACATCTCCACTCCCATCTTGCCTCAAGAAGCAATATAAA	858
QY	182	ValAlaValIleProIleAsnGlySerProArgThrProArgArgArgGlyGlnAsnArgSer	201
10	859	ACAGCTGCTTATACCCATAAATGGTTTACCTTCGCAACACCCAGCGAGGTTCAGAACAGGAGT	918
QY	202	AlaArgAlaValAlaGlySerGlnAsnAspArgArgIleIleGluValLeuCysGlySerGlu	221
10	919	GCAGCATATGCAAAATATAAAATATATATATAATTTGAATTTCTCTCTGTAAGAA	978
QY	222	HisGluCysAsnIleAspGluValIysAsnValTyrPheCysAsnPheIleTyrPheMet	241
10	979	CAUCAAGTAAATACATCAGGAGAAAAGTTTATTTCAAAAATTTATACCTTTTATC	1038
QY	242	AsnSerLeuGlyLeuValThrSerAsnGlyLeuProIleValGluAsnLeuSerLysArg	261
10	1039	AAATCTCTTCAGATCTGTAACATCTAATGGACTTCACAGGTTCGAAATCTTCTTAACCA	1098
QY	262	TyrGluGluIleTyrLeuLysAsnIysAspLeuAspAlaArgIleuPheIleuAspHisAsp	281
10	1099	TACAGAGAATTTATCTTAAATATAATATATAGATGCAGATATTTTTTGGATATCATGAT	1158
QY	282	LysThrLeuGluThrAspSerIleAspSerPheGluThrGluArgThrProArgLysSer	301
10	1159	AAAACCTCTTCAGACTATCTATACACAGTTTGAACACACAGACACACACCAACCAAAAGCT	1218
QY	302	AsnLeuAspGluGluValAsnValIleProProHisThrProValArgThrValMetAsn	321
10	1219	AACCTTGAACACAGCTGAATGATATCTCCACACATCCAGTTAGAGATGTATATCAAC	1278
QY	322	ThrIleHisGlnIleuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu	341
10	1279	ACTATGCAAAATTAATGATGATTTTAAATTCAGCAATGATCAACGCTTCAGAAATCTG	1338
QY	342	IleSerTyrPheAsnAsnCysThrValAsnProLysGluSerIleLeuLysArgValLys	361
10	1339	AATTCCTATTTTAACTTCGACGTGAATCCAAAGCAAAAGTATATCTCAAAAGAGTCAAG	1398
QY	362	AspIleLeuTyrIlePheLysGlnLysPheAlaLysAlaValGlyGlnGlyCysValGlu	381
10	1399	GAATATAGGATACATCTTAAACAGAAATTTCTCTAAAGCTGCGCACAGGCTGTCTCGAA	1458
QY	382	IleCysSerGlnArgTyrIleLysGlyValArgIleuTyrTyrArgValMetGluSerMet	401
10	1459	ATTGATATCAACAGATACAACTTGGACTTCGCTGTATTACCGAGTAATGGAATCCATG	1518
QY	402	LeuLysSerIleGluGluArgLysArgSerIleGlnAsnPheSerLysLeuLysAsnAspAsn	421
10	1519	CTTAAATCAGACACAGCAATACATCACTCAAAATTTACCAAACTTCGAGATGACAAC	1578
QY	422	IlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrTyrSerArg	441
10	1579	ATTTTCATATCTCTTATTTGGCGTGTCTTCAGCTTCTAATGGCCACATATAGCACA	1638
QY	442	SerThrSerGlnAsnLeuAspSerClyThrAspLeuSerPheThrGlnIleLeuAsnVal	461
10	1639	AGTATATTTCAATCTTATTTCTGTAACATATAATTTCTTCTTCATCGATCTCGAAGTG	1698
QY	462	LeuAsnIleLysAlaPheAspPheTyrLysValIleGluSerPheIleLysAlaGluGly	481
10	1699	CTTAAATTAAGAGCTTTCATTTTCAACAACTGATCCAAAGTTTATCAACGCACAGAGC	1758
QY	482	AsnLeuThrArgGluMetIleLysHisLeuGluArgCysGluHisArgIleMetGluSer	501
10	1759	AACCTTCAACAGCAATATCAATAACATTTAGAACGATCTCAACATCGAATTCGCAATCC	1818
QY	502	LeuAlaIleLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGlu	521
10	1819	CTTGATAGCTCTCAGATTCACCTTTATTTGATCTTATTATTAACATCAAGAGACCAGAA	1878
QY	522	ClyThrThrAspHisGluSerAlaCysProLeuAsnLeuLeuGlnAsnAsnHis	541

## RESULTS

AL307031  
AL307031

AR207031  
LOGIC $\frac{f_2}{f_1}$ 

\_\_\_\_\_

[illegible]

## RESULTS

AL307031  
AL307031

AR207031  
LOGIC $\frac{f_2}{f_1}$ 

\_\_\_\_\_

DEFINITION Sequence from patient US 6379927.  
 ACCESSION AR207843.1  
 VERSION AR207843.1  
 KEYWORDS unknown.  
 ORGANISM unknown.  
 REFERENCE 1 (bases 1 to 2994)  
 AUTHORS Antelman, D., Gregory, R.J., and Willis, K.N.  
 TITLE Notch1 and Notch2 fusion proteins  
 JOURNAL Patient: US 6379927, 3: 26 Apr 2002.  
 FEATURES  
 Location/Qualifiers  
 1..2994  
 /organism="unknown"

BASE COUNT 974 a 618 c 594 g 809 t  
 GC=31.81%

## Alignment Statistics:

Prod. No.: 9,846,412 Length: 2994  
 Score: 4494.50 Matches: 867  
 Percent Similarity: 94.54% Conservative: 1  
 Best Local Similarity: 94.44% Mismatches: 1  
 Query Match: 98.94% Indels: 59  
 Gaps: 1

US 09 026 459A 47 (1 865) x AR207843.1 (1-2994)

QY 1 MetProProGysThrProArLysThrAlaAlaAlaThrAlaAlaAlaAlaAlaGluPro 20  
 DB 149 ATGAGAGGCAAAACGCGGCAAAACGCGGCAAGCGCGCGCGCGCGCGCGCGCGCG 198  
 QY 21 ProAlaProProProProProProGluGluAspProGluGluAspSerGlyProGlu 40  
 DB 199 CCGGAGAGCG 258  
 QY 41 AspLeuProLeuValArgLeuGluPheGluGluThrGluGluAlaThrAlaLeu 60  
 DB 259 GAGTGGCTCTGTCAGGTGAGTTTCAGAGAACAGAACAGAACAGAACAGAACAGAAC 318  
 QY 61 CysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80  
 DB 319 TGTAGAAATTAAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 378  
 QY 81 ValSerSerValAspGlyValLeuGlyValLeuGlyValLeuGlyValLeuGlyVal 100  
 DB 379 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438  
 QY 101 ThrCysThrPheThrAlaAlaValAspLeuAspGluMetSerPheThrPheThrGluLeu 120  
 DB 439 ATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT 498  
 QY 121 GluGlyAsnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 140  
 DB 499 CAGAAAAAATAGAAAAATAGAAAAATAGAAAAATAGAAAAATAGAAAAATAGAAAAAT 558  
 QY 141 SerThrValAspAsnAlaMetSerArgLeuLeuLysLysTyrAspValLeuPheAla 160  
 DB 559 AGAGCAAG 618  
 QY 161 LeuPheSerLysLeuGluAlaThrCysGluGluLeuLeuLeuLeuLeuLeuLeuLeu 180  
 DB 619 CTCTTCT 678  
 QY 181 Met 181  
 DB 679 ATATCTATCTAATAATAATAATCTGATTGTGTCTAATAATTTCTTGGATCAATTTTATTA 738  
 QY 181 181  
 DB 739 GTTAAAGAGAGATATACAAAGAGAGATATACAAAGAGAGATATACAAAGAGAGAT 798  
 QY 181 181

DB 799 TGTGTCCTTGAATATTTTATTAATCTCTATCTCTAGTGGTCAAAATCAATAATAA 858  
 QY 182 ValAlaValThrProThrAsnGlySerThrArgThrProAlaAlaGlyGluAsnAlaSer 201  
 DB 859 ACAGTCTGTTATACCACTTAAAGGTTCACTCGGAATATCTACGTCAGCTTACAAATAGAG 918  
 QY 202 AlaArgThrAlaLysGluLeuGluAsnAspThrArgThrAlaThrValLeuValLysGlu 221  
 DB 919 GCACGGATAGCAAAACCAACATAGAAAATCACTACAAAGAAATATCTAAGCTCTCTCTAATA 978  
 QY 222 HisSerCysAsnSerAspGluValLysAsnValLysThrCysAsnThrThrProMet 241  
 DB 979 CATGAATGTATATAGATAGAGGTGAAAAATGTTTATTTAAAAATTTTATATTTTATAG 300  
 QY 242 AsnSerLeuGlyLeuValThrSerAsnGlyLeuThrGluValGluAsnLeuSerLysArg 261  
 DB 1039 AATCTCTTGGACCTGTAAACATCTAAAGGCTCTCAACAGTTCAAAAATCTCTTAAATCA 1098  
 QY 262 TyrGluGluThrLeuLysAsnLysAsnLeuAsnAlaAlaAlaLeuLeuGluGluAsp 281  
 DB 1099 TACGAAGAAAATTAATCTTAAAAATTAAGATCTACATCTAAAGATATATCTTGGATCA 340  
 QY 282 LysThrLeuGluThrAspSerThrAspSerPheGluThrGluAlaThrProAlaLysSer 401  
 DB 1159 AAAATCTCTACAGCTGATCTTATATAGACAGTTTGGAAATACAGAACATCTCTGAAAAAT 460  
 QY 302 AsnLeuAspGluGluValAsnValThrProProHisThrProValAlaThrValMetAsn 421  
 DB 1219 AACCTTCAACAGAGGAGCAATCTAAATCTCTCCACACATCTCTACGATAGCACTGAT 478  
 QY 322 ThrThrLeuGluLeuMetIleLeuAsnSerAlaSerAspThrProSerGluAsnLeu 441  
 DB 1279 ACTATCCAAATTAATCAATGATTTTAAATCTCAGCAAGTCAATCTCTGAAAAATCTG 500  
 QY 342 IleSerTyrPheAsnGlyThrValAspThrLysThrLysThrLysThrLysThrLys 361  
 DB 1339 ATTTCTCTATTTTACCACTGACAGTCAATCTGAAATCAAAATATATCTGAAAAATGTA 420  
 QY 362 AspThrLeuThrThrPheCysGluLysPheAlaLysValValGlyThrGlyValGlu 381  
 DB 1399 CATATAGATACATCTTAAAGCAAAATCTCTAAAGCTCTGACATAGGCTCTGCTGAA 440  
 QY 382 ThrThrSerGluAlaThrLysLeuGlyValAlaThrLeuThrTyrAlaValMetGluSerMet 401  
 DB 1459 ATTGCAATCAACGACATACAAATCTGACATCTGCTGCTGCTGCTGCTGCTGCTGCTG 460  
 QY 402 LeuLysSerGluGluGluThrArgSerThrThrAspSerThrLysThrLysThrLysThr 421  
 DB 1519 CTAAATCTAGAAAGAGAGAGATATCTTCAAAATTTTACAAATTTTACAAATTTTACAA 480  
 QY 422 ThrThrHisMetSerLeuLeuAlaCysAlaThrLeuValValMetAlaThrLysSerArg 441  
 DB 1579 ATTTCTCAATGCTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500  
 QY 442 SerThrSerGluAsnLeuAspSerGlyThrAspLeuSerPheProThrThrThrThrThr 461  
 DB 1639 ACTACATCTCAAACTCTGATCTGCAACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 520  
 QY 462 LeuAsnLeuLysAlaPheAspPheThrLysValThrGluSerThrThrThrThrThrThr 481  
 DB 1699 CTAAATTTAAAGCTCTTGTATTTTAAAGTATGCAAAATTTTATTAAGGATTAAGGAT 540  
 QY 482 AsnLeuThrArgGluMetThrLysHisLeuGluThrArgLysGluHisAlaThrMetGluSer 501  
 DB 1759 AACCTTCAACAGAGAGATATAGAAATCTTACAAATCTGCAACATGCAATCTGCAATCT 560  
 QY 502 LeuAlaThrLeuSerAspSerProLeuPheAspLeuLysThrLysThrLysThrLysThr 521  
 DB 1819 CTTCATGCT 580  
 QY 522 GlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuProLeuGluAsnAsnHis 541  
 DB 1879 GAGCAACATGATCATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600











[illegible]

OY	302	AsnLeuAspGluGluValAsnValIleProTPTORHisThrProValArgThrValMetAspLys
DB	1219	AACCTTGATCAAGAGGTAAAGTAATTCGCCAACAACCCAGTTAGAGCTGTATTAAAC
OY	322	ThrIleGlnGlnLeuMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu
DB	1279	ACTATCCCAACAAATTAATCAATGATTTTAAATTCAGAACCTGATCAACCTTCAGAAAAATCIG
OY	342	IleSerTyrPheAsnAsnCysThrValAsnProLysGlnSerIleLeuLysAlaValLys
DB	1339	ATTTCCTATTATAAATGAATGAAGAACAAAGAAAGAAATATGAAAANATGGAG
OY	362	AspIleGlyTyrIlePheLysGluLysPheAlaLysAlaValGlyGlnGlyCysValGlu
DB	1399	GATATAGCATATCTTTAAAGAAGAAATTTATTAAGGCTGTGGACAGAGTTHGTGAAA
OY	382	IleGlySerGlnAlaQTrLysLeuGlyValAsqLeuTyrTyrArgValMetGluSerMet
DB	1459	ATTGCATCACACCAYACAACCTTGAGAGTTCCTTCTGATATACCGAGIAATCGAAATCCATC
OY	402	LeuLysSerGluGluGluArgLeuSerIleGlnAsnPheSerLysLeuLeuAsnAspAsn
DB	1519	CCTAATCAGAAGAAGAACCATATCATCTCAAAATTTTAGCAAACCTCTCAATGACAAAC
OY	422	IlePheHisMetSerLeuleuAlaCysAlaIleCuGluValMetAlaThrTyrSerArg
DB	1579	ATTTTTATATATTTTTATAGAGTGGTCTTGAGAGTTGTAAATGATCATATATAGCA
OY	442	SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProTrpIleLeuAsnVal
DB	1639	ACTACATCTCAGAACTCTGATCTCGAACAGATCTTCCTCCCCAAGGATCTGCAATGIC
OY	462	LeuAsnLeuLysAlaPheAspPheTyrLysValIleGlnSerPheIleLysAlaGlnGly
DB	1699	CCTAATTTAAAGCGTTGTATTTTAAAGTGATCGAAGATTTTATCAAAGGAAAGAGG
OY	482	AsnLeuThrArgGluMetIleTyrHisIleuGlnArgCysGluHisArgIleMetGluSer
DB	1759	NACCTGACAAAGAAATCATAAAATATTAGAAATGAGAAATCATCATCATCATCATC
OY	502	LeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGlu
DB	1819	CTGCAAGGCCTCTCAGATCCACTTTATTTGATCTATTAAACAATCAAGACGACAGAA
OY	522	GlyProThrAspHisLeuGlnSerAlaCysProLeuAsnLeuProLeuGlnAsnAsnHis
DB	1879	GAACCAATCATATACCTTATGTTTGCTCTTAACTTCTCTCCAGAAATATCATC
OY	542	ThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLysCylsOrThrThr
DB	1939	ATGTGACAAATATGTATNTTCTGTATAGATATCAAGAAAGAAAGAGTTCAATAGG
OY	562	ArgValAsnSerThrAlaAsnAlaGlnThrGlnAlaThrSerAlaPheGlnGlnLeuLys
DB	1999	CTGTATAATCTACTGCAATATCCAGACACACAACTCACTCAGCTTCCAGACACAGAG
OY	582	ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyrAlaIleValTrpLeu
DB	2059	CAATTCAATATACCTCTCTTCACTGTTTTATAAAAAGTGTATGAGTCAAGTATCTC
OY	602	ArgCuAsnThrLeuCysGluArgGlnCuSerGlnHisProGlnCuLuisIleIleGln
DB	2119	CAATTAATAATCTTTGTAATGTTTTCTTGAGAACCTAGAAATAGAAATATATCATC
OY	622	TrpThrLeuPheGlnHisThrLeuGlnAsnGlnTyrGlnLeuMetArgAspAlaGlnLeu
DB	2179	TGCACCTCTTTCAGCACACCTCTCCAAAGACTATGAACCTCATGACAGACAACCAATC
OY	642	AspGlnIleMetMetCysSerMetTyrGlyIleCysLysValLysAsnIleAspLeuLys
DB	2239	GACAAAATATATATGTGTCTCATGTATGATATATGTAAGAGTGAATATATACCTTAA
OY	662	PheLysIleIleValThrAlaTyrHisAspLeuProGlnAlaValGlnGlnThrPheLys